Evaluating macro-finance interactions using mixed frequency methods

by

Annika Lindblad

Doctoral dissertation, to be presented for public examination
with the permission of the Faculty of Social Sciences
of the University of Helsinki,
in the Lecture room of Economicum, Arkadiankatu 7,
on the 29th of November, 2019 at 12 noon.

Research Reports
Publications of the Helsinki Center of Economic Research, No. 4/2019
Dissertationes Oeconomicae

Annika Lindblad

Evaluating macro-finance interactions using mixed frequency methods

ISBN: 978-952-10-8754-7 (print)
ISBN: 978-952-10-8755-4 (online)
ISSN: 2323-9786 (print)
ISSN: 2323-9794 (online)

Unigrafia Oy
Helsinki 2019

Abstract

This dissertation examines how macroeconomic variables influence financial market volatility and correlations using mixed frequency time series methods. The modelling framework allows combining high-frequency and low-frequency data within the same model and thus allows directly relating the economic data to the low-frequency component of volatility or correlations. The dissertation sheds light on which economic variables influence the low-frequency component of volatilities and correlations, as well as examines various methods to improve long horizon forecasts for stock market volatility by utilising the information in macroeconomic variables.

The first essay considers the relative and combined importance of macroeconomic fundamentals and survey-based sentiment data for modelling US equity market volatility in a GARCH-MIDAS framework. It uses a data set which accurately takes into account real-time data revisions to lags of the macroeconomic data and extends the analysis to include several new variables. Forward-looking macroeconomic data is important for forecasting volatility, even after the information in sentiment indicators is controlled for. On the other hand, for example, consumer confidence indicators contain information complementary to forward-looking macroeconomic variables. Overall, models combining macroeconomic and sentiment data tend to improve in-sample fit and in some cases also out-of-sample forecast accuracy compared to models only driven by one type of data. The improvements in forecasting performance are, however, not statistically significant, and therefore the results do not strongly advocate using several explanatory variables in the MIDAS polynomial.

In the second essay I assess the time-variation in predictive ability arising from the inclusion of macroeconomic and financial data in a GARCH-MIDAS model for US stock market volatility. I compare forecasts from a GARCH-MIDAS model to forecasts from a nested GARCH model, and therefore the differences in forecasting performance directly reflect the impact of economic data. While forecasting performance between the two models is similar when considered over the full out-of-sample period, there is clear time-variation in relative forecasting performance over sub-samples. I suggest the variation could arise from the phase of the business cycle or the volatility environment and find particularly strong evidence in favour of economic variables being important for volatility forecasting during low-volatility periods. Forecast combination methods and a decision rule based on conditional predictive ability produce consistently better forecasts than the GARCH model, although statistical significance of the improvements depend on the loss function considered.

The third essay considers the time-variation in the co-movement of equity returns

and exchange rate returns in several markets using the DCC-MIDAS model. Determining the economic drivers of the low-frequency correlation aids in differentiating between the various theoretical explanations for the correlation, which predict both a positive and a negative relationship. The essay concentrates on the portfolio rebalancing channel and on two hypotheses suggested in the earlier literature, namely flight-to-quality and quantitative easing (QE) related search-for-yield, in addition to examining the sensitivity of the correlation to other economic variables related to portfolio rebalancing motives, such as the business cycle. Although there are common elements driving the return correlation in the different markets, for instance, interest rate differentials and quantitative easing measures, their impact on the correlation varies, suggesting the underlying theoretical explanation differs across markets. While the onset of US QE1 had a clear impact on the correlations, overall the results suggest that being in a QE regime is more important than announcement effects for the long-term correlation.

Acknowledgments

This dissertation began with a vague idea that I would like to know more about the linkages between the macro economy and the financial markets. Quickly I managed to mix in mixed frequencies and forecasting, and here we are, six years later. It has been a privilege to be able to fully focus on research, and on deepening my understanding of issues and methodologies which interest me, for several years. What is certain, however, is that all the obstacles would have been much more difficult to overcome without all the encouraging people around me.

Several people have given me invaluable support during this journey. In particular, I am grateful to my supervisor, Professor Markku Lanne, who never seemed to lose hope in the project and who always managed to steer me back on the right path. His advice on everything from early research ideas to submitting to journals has been extremely valuable. However, even more importantly, already during my master's degree, over a decade ago, he inspired me into the field of macroeconometrics. I have not regretted my choice of research area.

I am grateful to Aarhus University and CREATES, where I spent a semester focusing fully on my research. There I got the opportunity to discuss my work with Professor Charlotte Christiansen, whose insights were important for improving this thesis. I also want to thank everyone who commented on my papers in seminars and conferences, and in particular Docent Henri Nyberg and Assistant Professor Matthijs Lof, who gave insightful feedback on preliminary research projects at the HECER (nowadays Helsinki GSE) econometrics workshops. In particular, I highly appreciate Henri's encouragement and genuine interest in my research throughout my PhD.

I am thankful to Professor Hossein Asgharian and Associate Professor Ai Jun Hou for agreeing to act as pre-examiners of this thesis and for their helpful comments regarding my work. I also want to thank Professor Christian Conrad for agreeing to act as my opponent.

I am grateful for the financial support I received from the OP Group Research Foundation and the Yrjö Jahnsson Foundation, which enabled me to focus on the PhD full-time. Presenting my research at international conferences and attending summer schools would not have been possible without the financial support of the Academy of Finland, the Foundation for the Advancement of Finnish Securities Markets, the Otto A. Malm Foundation, and the travel grants of the University of Helsinki.

The PhD would surely have felt much heavier without the abundant peer support offered by the strong community of PhD students at Economicum. In particular, thank you Jaakko, Juho, Kristine, Michaela and Olli for all the lunches, coffee breaks and other breaks which provided a much needed respite from the PhD. A huge special thank you goes to Julia, with whom I shared an office for almost the whole duration of my PhD. I deeply appreciate all the relevant and irrelevant conversations we had during the years, as they made life at Economicum so much more enjoyable.

Lastly, I want to take the opportunity to thank my family and friends for their unwavering support and encouragement. I could always count on them to remind me that there is life outside – and after – the PhD.

Helsinki, October 2019
Annika Lindblad

Contents

1 Introduction

1.1 Background

Understanding the linkages between the macro economy and the financial markets is important for both market participants and policy makers, and the area has attracted renewed research interest since the financial crisis. It is well established by, for example, Fama and French (1989) and Schwert (1989a), that risk premiums and stock market volatility are countercyclical, laying the foundation for studying the connection between financial markets and the macro economy. Macroeconomic conditions influence financial markets by, for instance, affecting the confidence of economic agents, impacting portfolio rebalancing motives of international investors and influencing the attractiveness of investment opportunities by, for example, impacting the expected dividend flow or the expected discount rates. Investment decisions are influenced by both the volatilities and the correlations of the financial markets, and it is, therefore, essential to understand which factors affect the volatility of a market as well as the co-movement of different markets. Thus, it is both interesting and important to consider closely the connections between macroeconomic conditions and financial markets.

In recent years there has been a growing interest in high-frequency studies as more high-frequency data has become available. Nevertheless, our ability to model and forecast volatility or correlations over very short time horizons has not removed or lessened the need to understand how financial markets behave over longer time spans, such as several months or even years. Longer horizon forecasts are important both for policy makers and investors, especially from a risk-management perspective. Understanding how the macro economy impacts financial markets aids in, for example, designing policies which limit adverse financial market outcomes or makes it possible to hedge against adverse outcomes further ahead in the future. From a longer-horizon modelling and forecasting perspective we might be interested in disentangling and forecasting a trend component of volatility or correlations. When this trend component can be linked to macroeconomic developments, it could be possible to use it to improve forecasts, in particular over long horizons. The short-term volatilities or correlations then fluctuate around this low-frequency trend component. The short-horizon, high-frequency approach and the longer term, macro economy based perspective thus complement each other.

This dissertation considers the importance of economic variables for modelling and forecasting financial market volatilities and correlations using modern time series methods, which utilise data at different frequencies within the same model. Focus is on low-frequency movements in volatility and correlations, which can be linked to economic fundamentals. In the first two essays I highlight the importance of forward-looking data for volatilities and conclude that during low volatility periods macroeconomic data is especially useful for forecasting over long horizons. In the last essay I establish that while similar economic variables drive the long-term correlation between foreign exchange returns and equity returns in several markets, their impact is market-specific and no single previously suggested theoretical explanation is able to alone account for the empirical fluctuations in the correlations.

1.2 The econometric framework

Volatility has been modelled using various GARCH specifications and predictive regressions in the previous literature. This dissertation utilises recently developed univariate and multivariate GARCH models augmented by a MIDAS (MIxed DAta Sampling) component, which allows low-frequency data to directly enter a model for high-frequency volatility or correlations. The univariate GARCH-MIDAS model for volatilities was introduced by Engle et al. (2013), while the multivariate DCC-MIDAS model for correlations was developed from the DCC model of Engle (2002) by Colacito et al. (2011) and further extended to include macroeconomic variables by Conrad et al. (2014). The MIDAS framework, which enables mixing data sampled at different frequencies in a parsimonious way, was introduced by Ghysels et al. (2004).

The research on the macroeconomic determinants of stock market volatility has its roots in Schwert (1989b) and Officer (1973), but much of the early literature found links that were weaker than expected. Component GARCH models – where volatility is decomposed into a transitory high-frequency component and a slowly evolving low-frequency component – have recently provided robust evidence in favour of macroeconomic determinants of (low-frequency) stock market volatility, as in, for example, Engle and Rangel (2008), Engle et al. (2013) and Conrad and Loch (2014). The multivariate DCC-MIDAS model has been used to model, for example, the oil-stock correlation (Conrad et al. (2014)), stock market correlations (for example, Asgharian et al. (2016)) and the stock-bond correlation (for example, Asgharian et al. (2015)).

A benefit of using a GARCH-MIDAS or DCC-MIDAS model is that aggregating data to the lowest common frequency, as is commonly done in, for example, predictive regressions, leads to higher-frequency information being lost. In the GARCH-MIDAS and DCC-MIDAS models the lower frequency information is incorporated directly into the higher frequency model through the MIDAS polynomial, and no aggregation of the data is necessary. In addition, a two-step approach, where a noisy proxy of volatility (or correlation) is attained in the first step and a regression framework is employed in a second step to explain the volatility (or correlation) leads to bias in the regression parameter estimates, as discussed in Engle et al. (2013). The one-step approach of the GARCH-MIDAS and DCC-MIDAS models, employed in this dissertation, allows the simultaneous estimation of both the low-frequency and the high-frequency components. The low-frequency data, which is measured at either quarterly or monthly frequency, is used to extract a long-term trend from daily financial market data. This allows relating the slowly moving trend component directly to economic variables, enabling an economic interpretation of the long-term trend.

1.2.1 Univariate GARCH-MIDAS framework

For estimating the volatilities I use the univariate GARCH-MIDAS model by Engle et al. (2013), where the returns on day t and in month τ can be modelled as having a multiplicative specification for the conditional variance:

$$r_{t,\tau} = \mu + \sqrt{m_{t,\tau}\, g_{t,\tau}}\, \varepsilon_{t,\tau}, \quad \varepsilon_{t,\tau} \mid \Phi_{t-1,\tau} \sim N(0,1), \quad \forall t = 1, ..., N_\tau \tag{1.1}$$

where $\Phi_{t-1,\tau}$ represents the information set up to day $t-1$, and N_τ is the number of trading days in period τ. $\sigma^2_{t,\tau} = m_\tau\, g_{t,\tau}$ is the total conditional variance, where m_τ[1] represents the (monthly) long-term volatility component and $g_{t,\tau}$ the (daily) GARCH component.

The daily $g_{t,\tau}$ component is modelled as a GARCH model, for example, an asymmetric GJR-GARCH(1,1) model (Glosten et al. (1993)):

$$g_{t,\tau} = 1 - \alpha - \beta - \gamma/2 + (\alpha + \gamma D_{t-1,\tau}) \frac{(r_{t-1,\tau} - \mu)^2}{m_\tau} + \beta g_{t-1,\tau}, \tag{1.2}$$

where $\alpha + \beta + \gamma/2 < 1$, $\alpha > 0$, $\beta \geq 0$ and $\alpha + \gamma \geq 0$. $D_{t-1,\tau}$ is an indicator function, taking the value 1 when $(r_{t-1,\tau} - \mu) < 0$ and 0 otherwise. Thus, the parameter γ describes the degree of asymmetry in volatility. When $\gamma = 0$ the standard GARCH(1,1) model is obtained.

The MIDAS polynomial with two explanatory variables takes the form:

$$log\ m_\tau = \bar{m}_v + \theta_1 \sum_{k=1}^{K_1} \varphi_k(\omega_{11}, \omega_{12}) X_{1,\tau-k} + \theta_2 \sum_{k=1}^{K_2} \varphi_k(\omega_{21}, \omega_{22}) X_{2,\tau-k}, \tag{1.3}$$

where K_i is the number of lags of explanatory data (X_i) included, and $\varphi_k(\omega_{11}, \omega_{12})$ and $\varphi_k(\omega_{21}, \omega_{22})$ are weighting schemes following a beta polynomial:

$$\varphi_k(\omega_1, \omega_2) = \frac{(\frac{k}{K})^{\omega_1 - 1}(1 - \frac{k}{K})^{\omega_2 - 1}}{\sum_{j=1}^{K} (\frac{j}{K})^{\omega_1 - 1}(1 - \frac{j}{K})^{\omega_2 - 1}}, \quad \text{where} \sum_{k=1}^{K} \varphi_k(\omega_1, \omega_2) = 1. \tag{1.4}$$

It is straightforward to see that the above model nests the GARCH-MIDAS-X model, where only one explanatory variable is included. If the variables do not affect stock market volatility (i.e., $\theta_1 = \theta_2 = 0$), all volatility is captured by the short-term component and the model collapses to the asymmetric GARCH model with $m_\tau = \exp \bar{m}_v$. The standard GARCH model is therefore nested in the GARCH-MIDAS model. The logarithmic specification ensures non-negativity of the long-term volatility component (m_τ) even when the explanatory variables take negative values. The sign of θ_i is interpretable: $\theta_i > 0$ ($\theta_i < 0$) implies that higher values of X_i are linked to higher (lower) long-term volatility.

[1] $m_{t,\tau}$ is fixed for all t in period τ. Hence, the subscript t is suppressed to ease notation and emphasise that m_τ evolves at a lower frequency than $g_{t,\tau}$.

The weighting scheme allows including a large number of the explanatory data in the model. The weight parameters, ω_1 and ω_2, govern the shape of the weighting scheme and can be freely estimated or fixed before estimation. The beta polynomial allows both monotonously decreasing weights ($\omega_1 = 1$) and hump-shaped weights ($\omega_1 < \omega_2$). If $\omega_1 = 1$ the rate of decay is determined by ω_2, where a larger value indicates faster decay. When ω_2 is very large all weight is on the first lag. If $\omega_2 < \omega_1$ all weight can be on distant lags. If $\omega_1 = \omega_2 = 1$ the weights are equal $(1/K)$ for all lags, which corresponds to a moving average.

1.2.2 Multivariate DCC-MIDAS framework

Following Engle (2009), the return vector follows the process $r_t \sim N(\mu, \mathbf{H_t})$. The conditional covariance matrix can be decomposed as $\mathbf{H_t} = \mathbf{D_t R_t D_t}$, where $\mathbf{R_t}$ is the conditional correlation matrix of the standardised return residuals, and $\mathbf{D_t}$ is a diagonal matrix with the standard deviations of the returns on the diagonal. Then, $\mathbf{R_t} = E_{t-1}[\xi_t \xi_t]$, where $\xi_t = \mathbf{D_t^{-1}}(r_t - \mu)$ are the standardised residuals obtained from the GARCH-MIDAS model. The short-term time-varying correlation for assets 1 and 2 are then estimated as:

$$q_{12,t} = \bar{\rho}_{12,\tau}(1 - a - b) + a(\xi_{1,t-1}\xi_{2,t-1}) + bq_{12,t-1}. \tag{1.5}$$

As noted by Engle (2009), $q_{12,t}$ can be thought of as an approximation of the true conditional correlation. Because the diagonal elements do not necessarily equal exactly one, a rescaling of the conditional correlation matrix is necessary:

$$\mathbf{R_t} = diag\{\mathbf{Q_t}\}^{-1/2}\mathbf{Q_t} diag\{\mathbf{Q_t}\}^{-1/2}.$$

The short-term correlations fluctuate around a long-term time-varying trend ($\bar{\rho}_{12,\tau}$). This long-term component can be specified in many different ways. The DCC-MIDAS-X model by Conrad et al. (2014) allows $\bar{\rho}_{12,\tau}$ to depend on economic data. To ensure the conditional covariance remains positive (regardless of the economic data) the Fisher-z transformation is used:

$$\bar{\rho}_{12,\tau} = \frac{exp(2\bar{z}_{12,\tau}) - 1}{exp(2\bar{z}_{12,\tau}) + 1}, \tag{1.6}$$

where

$$\bar{z}_{12,\tau} = \bar{m}_c + \theta_c \sum_{k=1}^{K_c} \varphi_k(\omega_c) X_{\tau-k}. \tag{1.7}$$

Here K_c is the number of lags of the explanatory data (X_τ) included in the long-term component and $\varphi_k(\omega_c)$ is a beta lag weighting scheme, as specified in equation (1.4), with $\omega_1 = 1$ fixed.

1.2.3 Likelihood function and estimation

The DCC-MIDAS model can be estimated by a two-step procedure, as described in Engle (2002) and Colacito et al. (2011). Following Engle (2002) the log-likelihood function can be written as:

$$LLF = -\frac{1}{2}\sum_{t=1}^{T}\left(n\ log(2\pi) + 2\ log|\mathbf{D_t}| + r_t'\mathbf{D_t^{-2}}r_t\right) - \frac{1}{2}\sum_{t=1}^{T}\left(log|\mathbf{R_t}| + \xi_t'\mathbf{R_t^{-1}}\xi_t + \xi_t'\xi_t\right)$$

From this expression it is clear why the model can be estimated in two steps: the first part of the likelihood function only depends on the data and parameters of the GARCH-MIDAS model, while the second part of the likelihood function depends on the standardised residuals and the DCC parameters. Therefore we can first use the first part of the likelihood function to estimate the variance parameters, and in a second step, conditional on the GARCH-MIDAS model parameter estimates, estimate the correlation parameters from the second part of the likelihood function.

The model can be estimated using quasi maximum likelihood estimation (QMLE). While consistency and asymptotic normality of the QML estimator for the rolling window GARCH-MIDAS model with realised volatility was established in Wang and Ghysels (2015), to my knowledge it has not been shown for the more general GARCH-MIDAS model with macroeconomic variables. As far as I know no estimation theory has been developed for the DCC-MIDAS model either yet.

1.3 Summary of the essays

This dissertation consists of three self-contained essays, which I will briefly summarise below. All three essays apply recently developed time series methods on data measured at different frequencies, with the aim of improving our understanding of financial market volatilities and correlations. The two first essays consider stock market volatility and its economic drivers, while the third essay examines the impact economic variables have on the correlation between equity market returns and currency market returns.

1.3.1 Chapter 2: Sentiment indicators versus macroeconomic data as drivers of long-term stock market volatility

The literature on the economic foundations of stock market volatility has found several useful predictors for volatility, starting from past realised volatility and including both macroeconomic and financial variables as well as principal components based on several variables. The GARCH-MIDAS literature has focused on studying this question by including one variable (or the level and volatility of the same variable) at a time in the MIDAS polynomial, with the exception of including realised volatility and one economic variable (as in, for example, Conrad and Loch (2014)) or summarising information in principal components (as in Asgharian et al. (2013)).

This essay, on the other hand, explores the relative and combined information content and predictive ability of macroeconomic fundamentals and survey-based sentiment indicators for US stock return volatility in the GARCH-MIDAS framework by Engle et al. (2013). In order to investigate how survey-based sentiment data affects the explanatory power of macroeconomic data, and vice versa, I include the two types of macroeconomic variables in the same MIDAS polynomial. The models are used both for in-sample analysis as well as out-of-sample forecasting. The goal is thus to better understand the sources of long-term volatility by considering the information content of survey-based sentiment indicators in relation to macroeconomic fundamentals.

I also introduce a number of novel variables for modelling and forecasting volatility in the GARCH-MIDAS context, which I believe could be indicative of volatility conditions, such as a comprehensive set of survey-based indicators, including forward-looking subcomponents of the consumer confidence index. I also argue that the recession probabilities given by professional forecasters proxy the expected business cycle, making them interesting predictors for stock return volatility. The new variables are mostly useful for modelling and forecasting volatility. In addition, I use, as far as possible, a real-time macroeconomic data set to match the information set of the agents at the time and accurately take into account real-time data revisions to lags. I find that using a data set with the lags revised in real time can lead to improved in-sample fit and slightly more accurate forecasts.

Regarding the relative importance of macroeconomic and survey-based sentiment indicators, I first of all find that once information in sentiment indicators is controlled for, backward-looking macroeconomic data contain only little additional useful information for modelling or forecasting stock return volatility. However, forward-looking macroeconomic variables remain useful for explaining stock market volatility even after sentiment data is included, in particular over long horizons. On the other hand, also some survey-based sentiment indicators improve forecasts compared to models only driven by forward-looking macroeconomic data.

Overall, the models combining the two types of data tend to improve the in-sample fit compared to models only driven by one type of data. There are also several cases in which the GARCH-MIDAS-X_1-X_2 models lead to improvements in out-of-sample forecast accuracy compared to both nested GARCH-MIDAS-X models, although the improvements are not simultaneously statistically significant, and therefore the results do not strongly advocate using several explanatory variables in the MIDAS polynomial.

1.3.2 Chapter 3: Evaluating the time-varying impact of economic data on the accuracy of stock market volatility forecasts

Forecasting volatility accurately is important, especially for financial market participants. While many papers compare forecast accuracy of volatility models over a long out-of-sample period, relatively little is known about the time-variation in forecasting performance arising from the inclusion of macroeconomic and financial data in a GARCH-MIDAS model for stock market volatility. When comparing forecasts from GARCH-MIDAS models to forecasts from a nested GARCH model the differences in forecasting performance directly reflect the impact of economic data. However, forecasting performance of a GARCH model and a GARCH-MIDAS model tends to be similar when compared over long out-of-sample periods (see, for example, Chapter 2 of this thesis). The main aim of this essay is to improve out-of-sample forecasts of long-horizon stock market volatility by considering how economic data contributes to volatility forecast accuracy over time and why forecast accuracy is time-varying.

This essay contributes to the current literature in three ways. First, I study the stability of the in-sample parameter estimates of GARCH-MIDAS models when estimated over rolling windows for US data and discuss their influence on in-sample fit and forecast accuracy. I also consider the impact of the weighting scheme of the predictors on both in-sample and out-of-sample results. This is crucial knowledge when estimating the GARCH-MIDAS model over sub-samples of the data and reveals how economic data is related to volatility. Second, I explore the additional time-varying predictive ability provided by macroeconomic and financial variables by comparing the out-of-sample forecasting performance of GARCH-MIDAS models to a GARCH model over subsamples. To consider potential reasons for the time-variation I investigate whether relative forecasting performance is affected by the business cycle or the market environment. Since the long-term trend component of volatility can be linked to macroeconomic variables, it is plausible that the forecasting ability of economic data depends on the economic or financial environment. Finally, I consider whether this information can be used to improve the accuracy of volatility forecasts and determine whether forecast accuracy can be enhanced by combining the individual GARCH-MIDAS model forecasts.

My main conclusion is that macroeconomic variables, as well as some financial variables, improve the accuracy of volatility forecasts significantly in low volatility periods and in particular over long horizons. Macroeconomic variables are useful to a lesser extent for forecasting volatility in recession periods as well, while financial variables otherwise struggle to perform even as well as the GARCH model. Especially when forecasting over long horizons there are shifts in forecasting performance over time, but time-varying forecast combination or selection methods mostly perform no better than simpler combination schemes. The term spread is the best variable for forecasting volatility over the 12 month horizon, especially in the early part of the sample. The forecast combination schemes and a decision rule based on conditional forecasting performance produce consistently better forecasts than the GARCH model on several horizons, but the statistical significance of the improvements depend on the loss function considered. As the GJR-GARCH model is rarely significantly better than the GARCH-MIDAS models and never significantly outperforms the combination forecasts, economic data can be considered useful for forecasting over long horizons. Finally, a rolling window estimation scheme leads to more volatile parameter estimates but better in-sample fit than an expanding window estimation scheme. Whether a rolling or expanding window is preferred for out-of-sample forecasting depends on the explanatory data, forecasting horizon and loss function.

1.3.3 Chapter 4: Economic origins of the dynamic co-movement of exchange rate and equity returns

Understanding the time-variation in the co-movement of various financial markets is important from a risk-management and investment perspective, as well as for policy makers. I begin this essay by noting that there is significant time-variation in the correlation between foreign exchange returns and equity returns for the US, the UK, Japan and the Euro area, in particular around the time of the financial crisis. Studying empirically the time-variation in the correlation is interesting and important, especially since there are conflicting theoretical models, some predicting a positive and others a negative relationship. The causal relationship between the equity market and the foreign exchange market has been studied extensively in the previous literature, but with mixed results. The economic drivers of the time-variation in the correlation between equity and currency returns has previously been considered in Moore and Wang (2014), who studied the impact of the trade balance and an interest rate differential on the correlation, and Kryzanowski et al. (2017), who used dummies related to the quantitative easing (QE) periods in the US to study the impact of the QE undertaken by the Federal Reserve on the correlation. Both papers first estimated the correlations using a DCC model and then used a regression model to relate the economic data to the correlation. They found a significant relationship between the correlation and the explanatory variables, but the effect varied across markets and variables.

In this essay I use the DCC-MIDAS model, introduced by Colacito et al. (2011) and further developed to include macroeconomic data in Conrad et al. (2014), to study the relationship between equity and exchange rate returns. I focus on variables related to international portfolio rebalancing and on determining how and why the sign of the long-term correlation changes over time. In particular, I consider two hypotheses suggested in the earlier literature, namely flight-to-quality (examined in Cho et al. (2016)) and QE related search-for-yield (considered in Kryzanowski et al. (2017)). In addition, I investigate whether the correlation is sensitive to standard macroeconomic variables related to portfolio rebalancing motives, such as the business cycle or the interest rate differential.

The results highlight the heterogeneity of the markets considered. Although there are common elements driving the return correlations, such as interest rate differentials and quantitative easing measures, their effect on the correlation varies and, therefore, so do the theoretical foundations of the dynamic correlations. For the US both the flight-to-quality and search-for-yield hypotheses are supported, as higher market volatility as well as more quantitative easing lead to an increasingly negative correlation. In the US particularly the first and second QE programmes have impacted the correlation strongly. Interest rate differentials are especially important in Japan, in line with the prominent role of the Yen as a funding currency. The results for Japan imply that the co-movement of FX and equities strengthens when the interest rate differential to the US shrinks and carry trade motives weaken. For the UK and the Euro area quantitative easing related variables are important, but the effect of QE on the correlation is positive Overall, the results suggest the balance sheet expansion itself is more important than QE announcements for the long-term correlation.

References

Asgharian, H., C. Christiansen, R. Gupta, and A. J. Hou (2016). Effects of economic policy uncertainty shocks on the long-run US-UK stock market correlation. CREATES *Research Paper 2016-29*.

Asgharian, H., C. Christiansen, and A. J. Hou (2015). Effects of macroeconomic uncertainty on the stock and bond markets. *Finance Research Letters 13*, 10–16.

Asgharian, H., A. J. Hou, and F. Javed (2013). The importance of the macroeconomic variables in forecasting stock return variance: a GARCH-MIDAS approach. *Journal of Forecasting 32*, 600–612.

Cho, J.-W., J. H. Choi, T. Kim, and W. Kim (2016). Flight-to-quality and correlation between currency and stock returns. *Journal of Banking & Finance 62*, 191–212.

Colacito, R., R. F. Engle, and E. Ghysels (2011). A component model for dynamic correlations. *Journal of Econometrics 164*, 45–59.

Conrad, C. and K. Loch (2014). Anticipating long-term stock market volatility. *Journal of Applied Econometrics 30*(7), 1090–1114.

Conrad, C., K. Loch, and D. Rittler (2014). On the macroeconomic determinants of long-term volatilities and correlations in US stock and crude oil markets. *Journal of Empirical Finance 29*, 26–40.

Engle, R. F. (2002). Dynamic conditional correlation - a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics 20*, 339–350.

Engle, R. F. (2009). *Anticipating correlations: a new paradigm for risk management*. Princeton: Princeton University Press.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *The Review of Economics and Statistics 95*(3), 776–797.

Engle, R. F. and J. G. Rangel (2008). The SPLINE-GARCH model for low frequency volatility and its global macroeconomic causes. *Review of Financial Studies 21*, 1187–1222.

Fama, E. F. and K. R. French (1989). Business conditions and expected returns on stock and bonds. *Journal of Financial Economics 25*, 23–49.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The MIDAS touch: Mixed data sampling regression models. *EFA 2009 Bergen Meetings Paper*.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Kryzanowski, L., J. Zhang, and R. Zhong (2017). Cross-financial-market correlations and quantitative easing. *Finance Research Letters 20*, 13–21.

Moore, T. and P. Wang (2014). Dynamic linkage between real exchange rates and stock prices: Evidence from developed and emerging asian markets. *International Review of Economics & Finance 29*, 1–11.

Officer, R. (1973). The variability of the market factor of the New York Stock Exchange. *Journal of Business 46*, 434–453.

Schwert, W. G. (1989a). Business cycles, financial crises, and stock volatility. *Carnegie-Rochester Conference Series on Public Policy 31*, 83–126.

Schwert, W. G. (1989b). Why does stock market volatility change over time? *Journal of Finance 44*, 1115–1153.

Wang, F. and E. Ghysels (2015). Econometric analysis of volatility component models. *Econometric Theory 31*, 362–393.

2 Sentiment indicators versus macroeconomic data as drivers of long-term stock market volatility[1]

2.1 Introduction

Stock market volatility is crucial for asset allocation and risk management, and it can also be interpreted as a measure of uncertainty. Therefore, it is important to understand, model and forecast stock return volatility accurately. It is well established by, for example, Fama and French (1989) and Schwert (1989a), that risk premiums and stock market volatility are countercyclical. The research on the macroeconomic determinants of stock market volatility has its roots in Schwert (1989b) and Officer (1973), but much of the early literature found links that are weaker than expected. The recent financial crisis underlined the importance of understanding the sources of volatility, leading to new interest in determining how the macroeconomy affects financial market volatility. Component GARCH models – where volatility is decomposed into a transitory high-frequency component and a slowly evolving low-frequency component – have recently provided robust evidence in favour of macroeconomic determinants of (low-frequency) financial market volatility.[2] Knowledge of the macroeconomic variables affecting volatility improves our understanding of why volatility varies over longer time periods and can enable more precise volatility forecasts, especially over long horizons.

The main contribution of this paper is to study the relative and combined information content and predictive ability of macroeconomic fundamentals and survey-based sentiment indicators for stock return volatility in the GARCH-MIDAS[3] framework of Engle et al. (2013). The GARCH-MIDAS model decomposes volatility into two components: a daily (GARCH) component which fluctuates around a long-term trend (for example, quarterly frequency). The low-frequency component is directly determined by economic variables. The GARCH-MIDAS literature has found a large set of useful predictors by including one variable at a time (or the level and volatility of the

[1]An early version of this essay is published as *HECER Discussion Paper No. 413.*

[2]For example, Engle and Rangel (2008), Engle et al. (2013) and Conrad and Loch (2014).

[3]GARCH-MIDAS stands for a generalised autoregressive conditional heteroskedasticity (GARCH) model combined with a mixed data sampling (MIDAS) approach. See Section 2.3.

same variable as in, for example, Engle et al. (2013)) in the MIDAS polynomial.[4] This paper explores whether survey-based sentiment indicators summarise the state of the economy and describe expectations of future macroeconomic conditions in a way that is either overlapping or complementary to macroeconomic fundamentals. In order to investigate how including survey-based sentiment data affects the explanatory power of macroeconomic data, and vice versa, I include the two different types of economic variables in the MIDAS polynomial simultaneously. The goal is thus to better understand the sources of long-term volatility by determining whether the in-sample fit and the out-of-sample forecasting ability of GARCH-MIDAS models can be improved by combining information in the two types of indicators. This paper thus builds on Engle et al. (2013), who introduced the GARCH-MIDAS model, and Conrad and Loch (2014), who, using the GARCH-MIDAS framework, found many macroeconomic and sentiment variables useful for modelling long-term stock market volatility.

First, I establish a baseline using GARCH-MIDAS models driven by one variable at a time. I use, as far as possible, a real-time macroeconomic data set to match the information sets of the agents at the time. This is to my knowledge the first paper to properly take into account real-time data revisions to lags in a GARCH-MIDAS model. In addition, I use a comprehensive set of survey-based indicators, concentrating on forward-looking subcomponents. I argue that the recession probabilities given by professional forecasters – a novel measure in this context – proxy the expected business cycle, making them interesting predictors for stock return volatility. Principal components are used to infer the usefulness of summarising information in many variables.[5]

Next, I determine the relative and combined importance of macroeconomic data and survey-based sentiment indicators as drivers of long-term stock return volatility by including different types of variables simultaneously in the GARCH-MIDAS model. These results are compared to the baseline obtained earlier. This allows us to infer the marginal benefit to the in-sample fit from adding a second variable into the model.

Finally, I explore the out-of-sample forecasting performance of the GARCH-MIDAS specifications in order to determine whether stock return volatility forecasts can be improved by simultaneously utilising information in macroeconomic variables and survey-based sentiment data. I take the GARCH(1,1) model as a benchmark for the out-of-sample forecasts, in line with Asgharian et al. (2013), but contrary to Engle et al. (2013) and Conrad and Loch (2014), who use a GARCH-MIDAS model where the long-term component is driven by realised volatility as a benchmark. A new perspective on the comparison of the baseline GARCH-MIDAS models is given by the Model Confidence Set (MCS) procedure by Hansen et al. (2011), which allows simultaneously comparing the performance of all models.

The main conclusions of this paper can be summarised in three points. First, the forward-looking components of the consumer confidence index generally outperform the main index, highlighting the importance of focusing on the forward-looking parts of survey-based sentiment indicators. In addition, realised volatility measured as the

[4]The exception is realised volatility, which was included together with macroeconomic data in the MIDAS polynomial in Conrad and Loch (2014) and Asgharian et al. (2013).

[5]Asgharian et al. (2013) also used principal components on a set of economic variables, which, however, differs clearly from the data set used in this paper.

sum of the absolute value of returns is more informative than the traditionally used sum of squared returns, both when it comes to modelling and forecasting stock market volatility. I also find that the one quarter ahead recession probability is useful for forecasting volatility over short horizons, while the four quarters ahead probability performs surprisingly poorly in the out-of-sample exercise. Using a data set with lags revised in real time can lead to a better in-sample fit and slightly more accurate forecasts than using only first-release data.

Second, once information in sentiment indicators is controlled for, backward-looking macroeconomic data, i.e., industrial production and the ADS index, contain only little additional useful information for modelling or forecasting stock return volatility. However, when the data set is augmented by professional forecasts for industrial production forecasting performance improves, especially over short horizons, compared to models using only sentiment indicators. Forward-looking macroeconomic variables, i.e., housing starts and the term spread, remain useful for forecasting volatility even after survey-based sentiment data is included, in particular over long horizons. The forward-looking nature of the explanatory data is overall crucial.

Thirdly, models combining the two types of data tend to improve the in-sample fit compared to the baseline models. There are also several cases in which the GARCH-MIDAS models driven by two variables lead to improvements in out-of-sample forecast accuracy compared to both baseline GARCH-MIDAS models, although the differences are not statistically significant and therefore the results do not strongly advocate using several explanatory variables in the MIDAS polynomial. The principal components perform generally well, but do not clearly outperform the best individual variables.

The remainder of the paper is organised as follows. Section 2.2 discusses the relationship between stock market volatility and the macro economy while reviewing the relevant literature. The GARCH-MIDAS framework of Engle et al. (2013) is presented in Section 2.3, and Section 2.4 describes the data. Section 2.5 presents in-sample results, while Section 2.6 discusses out-of-sample forecasts. Finally, Section 2.7 concludes.

2.2 Stock market volatility, survey-based sentiment data and the macro economy

While short-term volatility is well described and forecasted by GARCH models and stochastic volatility models[6], longer horizon modelling and forecasting of volatility relies on, for example, autoregressive models for realised volatility, predictive regressions and component GARCH models. It is widely accepted that stock return volatility is countercyclical, and that on the aggregate level the value of future cash flows depends on the state of the economy. The theoretical link between stock market volatility and the macro economy was formalised in, for example, Veronesi (1999), who presented a rational expectations equilibrium model where the stock market overreacts to bad news in good times and underreacts to good news in bad times. Other theoretical explanations include the present value models of Campbell (1991) and Campbell and

[6]See, for example, Poon and Granger (2003) for a survey, or Andersen et al. (2006) for an overview of volatility forecasting.

Shiller (1988), and models with time-varying volatility in fundamentals (for example, Bansal and Yaron (2004)). Mele (2007) developed a framework where countercyclical stock market volatility is a result of returns being more sensitive to changes in the economic environment when it is weak than when it is strong, resulting in risk premia being more volatile in bad times than in good times. It can be argued that stock market volatility affects the real economy, but also that the real economy affects stock market volatility. For example, the uncertainty hypothesis of Romer (1990) suggested that higher volatility on the stock market leads to higher uncertainty regarding future macroeconomic conditions, resulting in lower economic activity. On the other hand, a weaker economic environment leads to higher uncertainty regarding future investment opportunities, and hence increased uncertainty regarding the dividend flow, which can be reflected as higher stock market volatility.

The link from survey-based sentiment indicators to stock market volatility can be thought of as being directly analogous to the link between macroeconomic fundamentals and volatility: if sentiment indicators describe the current and/or expected economic situation, also confidence indicators should be linked to volatility in a countercyclical manner.[7] In particular, forward-looking sentiment data can plausibly relate to expectations of future dividends and returns. In the case of excess returns Campbell and Diebold (2009) used survey data to conclude that expectations regarding business conditions affect expected excess returns and reduce the explanatory power of more conventional financial predictors, such as the term premia.

In practice the role of sentiment data depends on whether economic agents (households, firms, analysts) form their expectations, summarised by sentiment indicators, on information already contained in macroeconomic fundamentals, or on a larger set of data also comprising information on, for example, expected economic conditions. In both cases confidence indicators might contain more information than macroeconomic fundamentals, which tend to be backward-looking or describe only one sector of the economy. If sentiment indicators contain additional information compared to macroeconomic fundamentals, they can be useful simultaneously when modelling and forecasting stock market volatility.

Empirically the success in linking macroeconomic variables to stock market volatility has been mixed. In his seminal paper Schwert (1989b) found that volatility is higher during recessions, but the evidence in favour of macroeconomic predictability of stock return volatility in the US is weak. The results echo those in Officer (1973). Extending the research of Schwert (1989b) to an international setting Davis and Kutan (2003) failed to establish a solid link between macroeconomic volatility and stock market volatility. Mixed results were reported by Errunza and Hogan (1998) (European and US data) and Pierdzioch et al. (2008) (German data), while Paye (2012), using US data and predictive regressions, found little evidence of out-of-sample predictability improvements using macroeconomic data over benchmark AR models, although forecast combinations help and Granger causality is found. On the other hand, Hamilton and Lin (1996) found that a bivariate ARCH framework with Markov-switching for industrial production and stock market volatility is useful for forecasting volatility in the US, with recessions

[7]This is in line with the "news" view of consumer confidence, i.e., that there is a relationship between confidence and the macro economy because confidence includes information regarding current and future states of the economy (Barsky and Sims, 2012).

accounting for a large part of variation in volatility. Using predictive regressions and a Bayesian Model Averaging approach Christiansen et al. (2012) showed that especially variables which can be thought of as proxies for credit risk, funding illiquidity or connected to the time-varying risk premia add significant out-of-sample predictive power for volatility in the US. Arnold and Vrugt (2008) showed that dispersion in the forecasts by professional forecasters is related to stock market volatility in the US, but the link disappears after 1996. For a large cross section of countries Diebold and Yilmaz (2008) determined that volatility in macroeconomic variables leads to more volatile stock markets.

Component GARCH models for stock return volatility, where the low-frequency component of volatility is driven by macroeconomic variables, have recently provided robust links between the macroeconomy and stock market volatility. Engle and Rangel (2008) suggested a Spline-GARCH model, which combines multiplicatively a high-frequency GARCH part and a slow-moving deterministic component based on macroeconomic variables. They found using a panel with 50 countries that macroeconomic volatility significantly influences low-frequency stock market volatility. Building on this idea Engle et al. (2013) developed the GARCH-MIDAS model, which combines a high-frequency GARCH component with a low-frequency component based on macroeconomic data and inspired by the MIxed DAta Sampling (MIDAS) literature, introduced by Ghysels et al. (2004). They found that macroeconomic data is useful for explaining and forecasting volatility in the US when performance is compared to a GARCH-MIDAS model with realised volatility driving the long-term component.

Using the GARCH-MIDAS framework and a wide selection of macroeconomic and sentiment variables Conrad and Loch (2014) concluded that macroeconomic data improves volatility forecasts in the US (compared to a similar benchmark as in Engle et al. (2013)) especially for long forecasting horizons. They focused on the lead-lag relationship between the variables and stock market volatility, emphasising the usefulness of including (professional) forecasts for backward-looking macroeconomic data in the MIDAS polynomial. Summarising information in US macroeconomic and financial data using principal components (PCs) Asgharian et al. (2013) concluded that a GARCH-MIDAS model driven by realised volatility and the first PC significantly improves the one-step-ahead forecast accuracy relative to a standard GARCH model when forecasting the long-term variance, while the improvements are not statistically significant when forecasting total stock market volatility. Asgharian et al. (2015) showed that macroeconomic uncertainty is a useful predictor of US stock market volatility.

Conrad and Schienle (2018) considered testing for an omitted long-term component in GARCH models. Their results suggested the one-component GARCH model is misspecified for stock market volatility, which motivates the use of a two-component model, such as the GARCH-MIDAS model. Conrad and Kleen (2019) provided further evidence in favour of multiplicative GARCH models, showing that models incorporating economic variables improve on the other volatility models for forecast horizons of two to three months. Amendola et al. (2019) introduced a GARCH-MIDAS model where the economic data is allowed to have an asymmetric effect on volatility and show that this model significantly outperforms standard GARCH models.

2.3 The GARCH-MIDAS model

The GARCH-MIDAS model by Engle et al. (2013) is a multiplicative two-component model for the conditional variance, where the high-frequency component is modelled as a standard GARCH model, while the low-frequency component is determined by economic data. The high-frequency component can be thought of as fluctuating around a slow-moving long-term trend driven by variables evolving at a lower frequency than returns. The MIxed DAta Sampling (MIDAS) approach[8] deals with the challenges related to using data sampled at different frequencies within the same model. A key feature of MIDAS is capturing the lag structure of the explanatory variables by a known function which depends on only a few parameters.

Following the interpretation in Engle and Rangel (2008), which builds on the log-linear dividend-ratio model in Campbell (1991) and Campbell and Shiller (1988), the stock return on day i and in period (month or quarter) t can be modelled as having a multiplicative specification for the conditional variance:

$$r_{i,t} = E_{i-1,t}(r_{i,t}) + \sqrt{\tau_{i,t}\, g_{i,t}}\, \varepsilon_{i,t}, \quad \varepsilon_{i,t} \mid \Phi_{i-1,t} \sim N(0,1), \quad \forall i = 1, ..., N_t$$

where $\Phi_{i-1,t}$ represents the information set up to day $i-1$ and N_t is the number of trading days in period t. $\sigma^2_{i,t} = \tau_{i,t}\ g_{i,t}$ is the conditional variance, where τ_t[9] is the long-term volatility component and $g_{i,t}$ the GARCH component. It is assumed that $E_{t-1}(r_{i,t}) = \mu$, that is, the expected return is constant. The model builds on the idea that the unexpected return, $r_{i,t} - E_{i-1,t}(r_{i,t})$, depends on news shocks, which affect dividends, interest rates or risk premia. The shocks can have short or long horizon effects, which motivates the division of volatility into a short-term and a long-term component.

It is well established that stock return volatility is asymmetric[10], i.e., that positive and negative news have different impact on volatility. Stock returns have been found to be negatively correlated with their volatility, and this has been attributed to the leverage effect (Black, 1976) or time-varying risk premia (see Awartani and Corradi (2005)). Hence, I use the asymmetric GJR-GARCH model by Glosten et al. (1993)[11]:

$$g_{i,t} = w + (\alpha + \gamma D_{i-1,t})\frac{(r_{i-1,t} - \mu)^2}{\tau_t} + \beta g_{i-1,t} \tag{2.1}$$

where $D_{i-1,t}$ is an indicator function, taking the value 1 when $(r_{i-1,t} - \mu) < 0$ and 0 otherwise. Thus, γ describes the degree of asymmetry in volatility. w is normalised to $w = 1 - \alpha - \beta - \gamma/2$ so that $E_{t-1}(g_{i,t}) = 1$. To ensure stationarity the condition $\alpha + \beta + \gamma/2 < 1$ is imposed. In addition, I assume $\alpha > 0$, $\beta \geq 0$ and $\alpha + \gamma \geq 0$ to ensure the variance remains positive.

Following Engle et al. (2013) the MIDAS polynomial with two explanatory variables

[8] Introduced by Ghysels et al. (2004) and discussed in detail in, for example, Ghysels et al. (2005), Ghysels et al. (2006), Ghysels et al. (2007), Andreou et al. (2010), and Wang and Ghysels (2015).

[9] $\tau_{i,t}$ is fixed for all i in period t. Hence, the subscript i is suppressed to ease notation and emphasise that τ_t evolves at a lower frequency than $g_{i,t}$.

[10] See, for example, Awartani and Corradi (2005) and the references therein.

[11] The same short-term component is used in Conrad and Loch (2014).

(X_1 and X_2) takes the form:[12]

$$log\ \tau_t = m + \theta_1 \sum_{k=1}^{K} \varphi_k(\omega_{11}, \omega_{12}) X_{1,t-k} + \theta_2 \sum_{k=1}^{K} \varphi_k(\omega_{21}, \omega_{22}) X_{2,t-k}, \tag{2.2}$$

where $\varphi_k(\omega_{11}, \omega_{12})$ and $\varphi_k(\omega_{21}, \omega_{22})$ are weighting schemes (see Figure 2.1 for examples), and K is the number of lags of explanatory data included. When X_1 is realised volatility Conrad and Loch (2014) name the model the GARCH-MIDAS-RV-X model. Following this convention the above model is called the GARCH-MIDAS-X_1-X_2 model. The logarithmic specification ensures non-negativity of the long-term volatility component (τ_t) even when the explanatory variables take negative values. If the variables do not affect stock market volatility (i.e., $\theta_1 = \theta_2 = 0$), all volatility is captured by the short-term component and the model collapses to the asymmetric GARCH model with $\tau_t = exp m$. The standard GARCH model is therefore nested in the GARCH-MIDAS model. The sign of θ_i is interpretable: $\theta_i > 0$ implies that higher values of X_i are linked to higher long-term volatility, and vice versa.

Conrad and Loch (2014) used the MIDAS polynomial in (2.2) to investigate whether economic variables (X_2) are important for volatility after information in past squared returns (X_1) have been accounted for.[13] In addition, Engle et al. (2013) studied the combined effect of the level and volatility of a macroeconomic variable. I concentrate on specifications including a macroeconomic (X_1) and a survey-based sentiment (X_2) variable, but also include specifications with realised volatility as a robustness check (see Appendix 2.C). The MIDAS polynomial thus allows directly comparing the importance of different types of variables within the same model.

Conrad and Loch (2014) showed that especially for backward-looking data, such as industrial production growth, feasible two-sided filters are useful, where forecasts of a variable (from, for example, the Survey of Professional Forecasters (SPF)) are included in the MIDAS polynomial:

$$log\ \tau_t = m + \theta_1 \sum_{k=1}^{K_{lag}} \varphi_k(\omega_{11}, \omega_{12}) X_{1,t-k} + \theta_1 \sum_{k=-K_{lead}}^{0} \varphi_k(\omega_{11}, \omega_{12}) X^{SPF}_{1,t-k|t-1}. \tag{2.3}$$

This specification thus combines a macroeconomic variable and its survey-based forecast. I will use the feasible two-sided filter for industrial production and housing starts, for which SPF forecasts exist, in the empirical analysis.

A flexible but parsimonious weighting scheme, commonly used in GARCH-MIDAS models, is the beta lag polynomial[14], which ensures positive weights (this ensures non-negativity of volatility) adding up to one (this normalisation allows identifying θ_1 and θ_2):

$$\varphi_k(\omega_1, \omega_2) = \frac{(\frac{k}{K})^{\omega_1 - 1}(1-\frac{k}{K})^{\omega_2 - 1}}{\sum_{j=1}^{K}(\frac{j}{K})^{\omega_1 - 1}(1-\frac{j}{K})^{\omega_2 - 1}}, \quad \text{where } \sum_{k=1}^{K} \varphi_k(\omega_1, \omega_2) = 1.$$

[12] Additional variables can be included in the MIDAS polynomial in a straightforward manner, but each variable increases the parameter space by three new parameters.

[13] Asgharian et al. (2013) considered the same question using a mostly predetermined weighting scheme.

[14] Weighting schemes are discussed in more detail in Ghysels et al. (2007).

The weight parameters, ω_1 and ω_2, govern the shape of the weighting scheme and can be freely estimated or fixed before estimation. The beta polynomial allows both monotonously decreasing weights ($\omega_1 = 1$) and hump-shaped weights ($\omega_1 < \omega_2$). If $\omega_1 = 1$ the rate of decay is determined by ω_2, where a larger value indicates faster decay. When ω_2 is very large (for example, $\omega_2 > 100$) all weight is on the first lag. If $\omega_2 < \omega_1$ all weight can be on distant lags. If $\omega_1 = \omega_2 = 1$ the weights are equal ($1/K$) for all lags, which corresponds to a moving average. Each explanatory variable has its own weighting scheme, meaning that the shape of the weighting scheme can be different for different variables included in the same MIDAS polynomial.

To assess how much the variation in a particular variable explains of the overall expected volatility Engle et al. (2013) suggested calculating variance ratios: $\frac{Var(log(\tau_t))}{Var(log(\tau_t g_{i,t}))}$. The variance ratio can be interpreted as a measure of in-sample fit in the sense that the higher the variance ratio is, the larger is the share of the total expected volatility that can be explained by the long-term component. However, a low variance ratio does not necessarily imply poor model fit as it can also be a result of smooth movements in the underlying variable (Conrad and Loch, 2014). The GARCH-MIDAS model can be estimated using maximum likelihood (QML if the assumption of normally distributed errors does not hold).[15]

2.4 Data

I use the continuously compounded daily stock market return on the CRSP index from January 1973 to September 2017. Due to data availability I concentrate on the quarterly frequency for the explanatory data with a sample period from Q1 1970 to Q3 2017.[16]

A natural explanatory variable for stock market volatility is (lagged) realised volatility. The sum of squared returns ($\sum_{i=1}^{N_t} r_{i,t}^2$) is a commonly used measure for realised volatility (for example, Engle et al. (2013)). However, already Taylor (1986) and Ding et al. (1993), among others, explored the advantages of using the absolute value of returns when modelling especially the low-frequency component of volatility. More recently, using MIDAS regressions and intra-daily data, Ghysels et al. (2006) concluded that absolute returns outperform squared returns when forecasting quadratic variation in returns on short horizons (up to one month).[17] To determine whether the absolute value of returns outperform squared returns on longer horizons I also use $\sum_{i=1}^{N_t} |r_{i,t}|$ as a measure of realised volatility.

As **macroeconomic data** I use industrial production, the Aruoba-Diebold-Scotti Business Conditions Index (ADS index) and housing starts. Industrial production is a traditional macroeconomic variable for modelling and predicting stock return volatility, as it is a timely measure of output in the economy. The ADS index, which includes,

[15]While consistency and asymptotic normality of the QML estimator for the rolling window GARCH-MIDAS model with realised volatility was established in Wang and Ghysels (2015), to my knowledge it has not been shown for the more general GARCH-MIDAS-X model.

[16]Three years of explanatory data is needed to estimate the model for the first period. The long sample period leads to the exclusion of, for example, the Economic Policy Uncertainty index by Baker et al. (2016).

[17]Ghysels et al. (2006) found that realised power, based on intra-daily data, is the best measure for realised volatility, but I focus on daily data which is widely available for a long time period.

for example, labour market indicators and industrial production, tracks business conditions in real time and covers a range of lagging economic indicators. Housing starts gives an early indication of future economic activity and can be considered a leading indicator. For both industrial production and housing starts I use the annualised quarterly rate of growth (i.e., 100 $((X_t/X_{t-1})^4 - 1)$). As the long-term component of the GARCH-MIDAS model depends on several lags of the explanatory data, taking into account revisions can be important. Hence, for the macroeconomic data I use the last available vintage in each quarter of the real-time data sets from the Philadelphia Fed.[18] Thus the value of the macroeconomic variable included for period $t-1$ at time t tends to differ (because of revisions) from its value for period $t-1$ at time $t+1$.[19] I also include the term spread, defined as the difference between the 10-year Treasury bond yield and the 3-month T-bill rate. The term spread has been a successful driver of stock return volatility (see Conrad and Loch (2014)), and it is an accurate predictor of future economic activity (see, for example, Bauer and Mertens (2018)).[20]

I define **sentiment data** as survey-based confidence indicators[21]:

- **Household sentiment:** University of Michigan consumer confidence data (first differences), including forward-looking subindices: the News Heard index and the Buying conditions index. The News Heard index can be seen as a proxy for general sentiment in the economy since it surveys the kind of news regarding business conditions the respondents have recently read. The Buying conditions index is chosen over other forward-looking subindices because it has the lowest correlation with the main index.[22]

- **Business confidence:** The forward-looking ISM New Orders index as well as the ISM Recession indicator (New Orders - Inventories) (levels). These describe the demand of manufacturing businesses, which can be seen as a proxy for near-term business conditions. Note that the survey asks about the changes in production, new orders etc. that occurred during the month. Thus, the ISM report does not survey expectations but rather gives a "real-time" assessment of the near-term economic situation.[23]

- **Professional expectations:** Survey of Professional Forecasters (SPF) data. To describe expectations regarding the business cycle I use the probability given by professional forecasters that GDP will decline one or four quarters ahead.[24] An interesting feature of the one quarter ahead recession probability is that it replicates, in real time, relatively well the official NBER recession dates. Hence it

[18] Prior to 2008 real-time vintages of the ADS index are unavailable.

[19] Appendix 2.B compares results for first release data and real-time revisions. Data that is revised in real time mostly leads to better in-sample fit and slightly smaller forecast errors.

[20] On the other hand, an inflation measure is not included because it did not perform particularly well in Conrad and Loch (2014).

[21] Thus the term spread, which can be argued to be a sentiment measure for the financial markets, is primarily labelled as macroeconomic data.

[22] The correlation of the Buying conditions index with the main index is 0.71 over the sample versus roughly 0.95 for the Expected index and the 12 months ahead Business conditions index.

[23] See `https://www.instituteforsupplymanagement.org/` for details.

[24] Two data points are missing from the early part of the four quarters ahead recession probability series. I replace these by values from the previous quarter.

can be seen as a valid proxy for the current economic situation, whereas the four quarters ahead probability can be argued to summarise expectations regarding the business cycle. In addition, for industrial production growth and housing starts I include the median forecasts of the SPF for the two-sided filters.[25]

Orthogonalising the sentiment data with respect to macroeconomic fundamentals would provide a more exact measure of pure sentiment, but since the aim of this paper is to disentangle the usefulness of variables included in previous research this is left for future work. Descriptive statistics, data sources and graphs of the data can be found in Appendix 2.A. Standard unit root tests confirm the stationarity of the data.

As I will include different types of data in the same MIDAS polynomial, I consider the information overlap in sentiment indicators and macroeconomic data using correlations. Table 2.1 shows that squared and absolute returns are, as expected, highly correlated (0.91). The correlation between the absolute value of returns and the economic variables varies between the virtually zero correlation with forward-looking variables, such as the term spread (0.03), and the relatively high correlation with coincident or lagging indicators, such as the ADS index (-0.43). The correlation between the recession probabilities one and four quarters ahead is small (0.09), while for the other sentiment measures the subindices are highly correlated with each other. Industrial production and the ADS index are highly correlated with the ISM indices and the one quarter ahead recession probability but only moderately correlated with the consumer sentiment indicators and not at all correlated with the four quarters ahead recession probability. Housing starts is weakly correlated with most of the sentiment indicators, while the term spread has a relatively high correlation with the ISM Recession indicator (0.46) but is only moderately correlated with the other sentiment measures. As expected, the one quarter ahead recession probability is highly correlated with contemporaneous measures for economic activity, while the four quarters ahead probability is mostly correlated with forward-looking variables.

I use principal components analysis to aggregate information in all macroeconomic and sentiment variables. As the variables are measured on different scales I base the principal components (PCs) on the correlation matrix. Table 2.2 shows the correlations between the eleven principal components and the eleven variables. The first PC is highly correlated with most of the variables, in particular with the ISM indices, the ADS index, industrial production and the one quarter ahead recession probability. Hence it captures current business conditions. The second PC has the highest correlations with the consumer confidence indices. The third PC is mostly correlated with the term spread and the four quarters ahead recession probability but also housing starts. Thus it describes the forward-looking components of the data. The remaining principal components are either primarily correlated with just one or two variables or not very correlated with any of the variables. Hence, I use the first three principal components as explanatory variables in the MIDAS polynomial. The first three PCs are also the ones which explain over 10% of the total variation in the data.

[25] As concluded in Conrad and Loch (2014), for housing starts, which is in itself a forward-looking variable, the gains of using a two-sided filter are limited. Thus, the results for the two-sided filter for housing starts and one-sided filter for industrial production growth can be found in Appendix 2.E.

Table 2.1: Correlation matrix for macroeconomic and sentiment data

	Σr^2	$\Sigma\|r\|$	CC	NH	BC	NO	RI	IP	HS	TS	ADS	1Q
$\Sigma\|r\|$	0.91	1										
CC	-0.10	-0.15	1									
NH	-0.10	-0.14	0.77	1								
BC	-0.15	-0.20	0.71	0.54	1							
NO	-0.31	-0.37	0.22	0.19	0.32	1						
RI	-0.23	-0.25	0.38	0.40	0.40	0.66	1					
IP	-0.30	-0.33	0.19	0.14	0.27	0.80	0.54	1				
HS	-0.20	-0.25	0.07	0.02	0.12	0.33	0.12	0.27	1			
TS	0.04	-0.03	0.19	0.17	0.22	0.26	0.46	0.09	0.09	1		
ADS	-0.39	-0.43	0.27	0.18	0.33	0.82	0.56	0.93	0.32	0.09	1	
1Q	0.33	0.40	-0.18	-0.16	-0.30	-0.73	-0.53	-0.66	-0.32	-0.23	-0.74	1
4Q	0.09	0.06	-0.19	-0.20	-0.20	-0.13	-0.22	-0.01	-0.05	-0.24	0.02	0.09

The variables used are: sum of squared returns (Σr^2), sum of absolute value of returns ($\Sigma|r|$), consumer confidence (CC), News Heard index (NH), Buying conditions index (BC), ISM New Orders index (NO), ISM Recession indicator (RI), industrial production growth (IP), housing starts (HS), term spread (TS), ADS index (ADS), SPF 1Q ahead (1Q) and SPF 4Q ahead (4Q) recession probabilities. Sample period: Q1 1970 - Q3 2017. Latest available data is used for the ADS index, housing starts and industrial production.

Table 2.2: Correlation matrix for the principal components and economic data

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11
CC	-0.59	-0.65	0.38	-0.04	-0.04	0.03	-0.09	0.06	0.14	0.27	0.04
NH	-0.52	-0.65	0.31	-0.02	0.00	0.35	0.18	-0.10	-0.06	-0.19	-0.02
BC	-0.63	-0.49	0.27	0.00	0.02	-0.48	-0.18	0.13	-0.02	-0.13	-0.01
NO	-0.83	0.39	-0.01	-0.02	0.19	0.04	-0.06	-0.01	-0.34	0.06	0.02
RI	-0.80	-0.03	-0.21	-0.21	0.05	0.29	-0.20	0.36	0.10	0.02	-0.00
1Q	0.76	-0.37	0.03	0.01	-0.08	0.17	-0.47	-0.16	-0.07	0.01	-0.02
4Q	0.25	0.38	0.53	-0.56	-0.45	-0.00	0.01	0.04	-0.04	-0.00	0.01
IP	-0.78	0.50	0.13	0.12	-0.02	0.05	-0.15	-0.22	0.14	-0.09	0.13
ADS	-0.83	0.45	0.15	0.12	-0.11	-0.01	-0.05	-0.14	0.07	0.04	-0.17
HS	-0.46	-0.25	-0.43	0.27	-0.68	-0.01	0.03	0.02	-0.06	0.02	0.03
TS	-0.40	-0.24	-0.63	-0.57	0.06	-0.12	0.02	-0.23	0.02	-0.02	-0.01
V %	42.27	18.91	11.36	7.13	6.61	4.41	3.29	2.79	1.50	1.28	0.45

The last row (V %) reports how much of the total variation in the data is accounted for by each principal component. Otherwise, see notes on Table 2.1.

2.5 In-sample results

In Section 2.5.1 I establish baseline results using the GARCH-MIDAS model with one explanatory variable (GARCH-MIDAS-X).[26] This section largely confirms the results in Conrad and Loch (2014) using a different stock return index and real-time macroeconomic data. Section 2.5.2 determines the relative and combined importance of macroe-

[26]The estimations are executed in Matlab, building on the basic code by Engle et al. (2013).

conomic variables and survey-based sentiment data by including two explanatory variables in the same MIDAS polynomial (GARCH-MIDAS-X_1-X_2).

2.5.1 Baseline results

The lag length (K) for the explanatory data in the long-term component is selected based on the data. I choose the K which maximises the value of the log-likelihood function when K is allowed to be 4, 8, 12 or 16 quarters. The value of the log-likelihood function tends to be maximised at 8 or 12 lags, levelling off after this. Therefore I use three years of lagged data, i.e., $K = 12$ (quarterly data). However, for the models driven by the News Heard index and the second principal component 16 lags are used, as this leads to a weighting scheme where the weights are close to zero for the last lag, the in-sample fit improves and the log-likelihood function values increase slightly.[27]

Next, whether the weighting scheme ($\varphi_k(\omega_1, \omega_2)$) is restricted ($\omega_1 = 1$) before estimation or not is determined for each explanatory variable based on a likelihood ratio test (LRT) between the two specifications. The choice is clear for all models except the ones driven by the absolute value of returns and the term spread. For the term spread I use the unrestricted weighting scheme as it gives more reasonable weights, a better in-sample fit and the term spread is usually considered a forward-looking variable (see, for example, Conrad and Loch (2014)). For realised volatility a restricted weighting scheme has been chosen in the earlier literature, and I follow this convention. As explained in Section 2.3, the restricted scheme forces the weights to be decaying, i.e., recent data matters the most for long-term volatility.

The GARCH model parameters are consistently, robustly and similarly estimated for all specifications (Table 2.3).[28] The parameter determining the degree of asymmetry in volatility (γ) is always highly significant and positive, indicating, as expected, that lower-than-expected returns lead to a higher conditional variance. The choice of an asymmetric GARCH model is thus well-motivated. In all specifications $\alpha + \beta + \gamma/2$ is clearly below one, indicating that the GARCH model is stationary. Overall, the GARCH parameters in the GARCH-MIDAS specifications get values roughly in line with the estimates for the basic asymmetric GARCH(1,1) model (last row in Table 2.3). Appendix 2.D presents graphs relating to the baseline results, showing a decomposition of volatility into its two components and the weighting schemes.

Parameter θ determines how the explanatory data affects long-term volatility. It is highly significant in all specifications and has the expected sign (Table 2.3): positive for realised volatility measures and recession probabilities, and negative for macroeconomic variables and sentiment indicators.[29] Positive (and highly significant) estimates for the recession probabilities indicate that a higher probability of recession among

[27] Since by construction the last weight in the beta polynomial is zero, I use $K + 1$ lags in the estimation, where the 13^{th} (or 17^{th}) lag gets zero weight. This follows Conrad and Loch (2014).

[28] I do not report the GARCH parameter estimates for the rest of the paper. They are similarly and robustly estimated throughout the specifications. Full results are available upon request.

[29] Notice that when testing the significance of θ, θ and the weight parameters ω_i are not separately identified under the null hypothesis, which affects the asymptotic distribution of the test statistic. However, I follow the convention in the GARCH-MIDAS literature (for example, Engle et al. (2013) and Conrad and Loch (2014)) and proceed using the standard t-statistic. See Ghysels et al. (2007) for a discussion of the problem in MIDAS regressions.

professional forecasters translates into higher stock market volatility. The estimates of θ for all three principal components are positive and therefore largely in line with the correlations between the factors and the explanatory variables (Table 2.2). Overall the results strongly support the countercyclical nature of long-term stock market volatility.

The variance ratio of the GARCH-MIDAS model where the long-term component is driven by the absolute value of returns is roughly 30%, which is clearly greater than for any other variable, including squared returns (18%). Realised volatility based on the absolute value of returns thus seems to incorporate a large amount of useful information for explaining long-term stock market volatility. The good in-sample fit is also evident from Figure 2.5b. Considering the macroeconomic and sentiment data, the long-term components driven by housing starts and the Buying conditions index explain a large share of the total variance (around 17%), while the variance of the long-term component determined by industrial production only accounts for 6% of the total variance. Using the feasible two-sided filter for industrial production improves the fit to almost 10%, but this is still the worst in-sample fit among all the models. The principal components driven models have relatively high variance ratios between 14.7% and 18.1%, indicating it can be useful to summarise information. Based on the variance ratios, the forward-looking News Heard index and the Buying conditions index outperform the main consumer confidence index, highlighting the gains of using forward-looking sentiment indicators.

The weighting schemes are in line with the lead-lag relationships established in Conrad and Loch (2014), who interpreted hump-shaped weighting schemes as a sign that the variable is forward-looking, while monotonously declining weights imply that the variable is lagging or coincidental (the weighting schemes are drawn in Figure 2.6). For the four quarters ahead recession probability the highest weight is on a relatively distant lag, while the fastest decay in weights is seen for the one quarter ahead recession probability, for which only the first five lags get a non-zero weight. This is intuitive, since one would not expect short-term recession probabilities from several years ago to influence volatility. The two-sided filter for industrial production growth puts a significant amount of weight on the leads, while for housing starts the weighting schemes are similar. The first principal component has decaying weights, while the second and the third PC have hump-shaped weighting schemes. These weighting schemes seem plausible, as the first PC is mainly correlated with indicators for current business conditions which have decaying weights themselves, while the second and third PC are mainly related to variables with hump-shaped weighting schemes.

The results remain robust to including realised volatility in the specifications (see Appendix 2.C) in line with the results in Conrad and Loch (2014). Thus, macroeconomic data and survey-based sentiment indicators explain parts of the long-term stock market volatility not captured by realised volatility.

Table 2.3: Estimation results of the GARCH-MIDAS-X model

Variable (X)	μ	α	β	γ	θ	ω_1	ω_2	m	LLF	VR
Sum of squared returns	0.0479***	0.0198***	0.8764***	0.1291***	0.0041***	1	7.2518**	-0.4800***	-14301.56	18.28
	(0.0072)	(0.0052)	(0.0143)	(0.0192)	(0.0005)		(2.1369)	(0.0804)	(0.5248)	
Sum of absolute value of returns	0.0482***	0.0145***	0.8627***	0.1391***	0.0206***	1	11.1342***	-1.1520***	-14276.34	30.66
	(0.0072)	(0.0053)	(0.0146)	(0.0191)	(0.0017)		(2.1544)	(0.1027)	(0.0869)	
Consumer confidence index	0.0471***	0.0198***	0.8942***	0.1154***	-0.1676***	1.7207*	2.5824**	-0.1773*	-14304.70	12.12
	(0.0073)	(0.0052)	(0.0145)	(0.0189)	(0.0302)	(0.3886)	(0.6414)	(0.0873)	(0.0050)	
News Heard index	0.0473***	0.0204***	0.8919***	0.1161***	-0.0886***	2.1897***	2.6234***	-0.1861**	-14299.39	15.97
	(0.0073)	(0.0052)	(0.0147)	(0.0189)	(0.0159)	(0.3381)	(0.4731)	(0.0835)	(0.0000)	
Buying conditions index	0.0472***	0.0172***	0.8887***	0.1222***	-0.1264***	1	1.9902***	-0.1782**	-14292.82	17.90
	(0.0073)	(0.0054)	(0.0143)	(0.0188)	(0.0186)		(0.2562)	(0.0780)	(0.1720)	
ISM New Orders index	0.0456***	0.052***	0.8990***	0.1173***	-0.0481***	1	4.9613**	2.4378***	-14303.01	13.36
	(0.0073)	(0.0054)	(0.0139)	(0.0184)	(0.0092)		(1.8739)	(0.5087)	(0.6105)	
ISM Recession indicator	0.0466***	0.0191***	0.8972***	0.1132***	-0.0680***	1	2.2415***	0.3804***	-14306.67	10.46
	(0.0073)	(0.0052)	(0.0139)	(0.0183)	(0.0120)		(0.4271)	(0.1315)	(0.1682)	
SPF 1Q ahead recession probability	0.0462***	0.0172***	0.8925***	0.1212***	0.0167***	1	15.7557	-0.5171***	-14300.06	12.53
	(0.0073)	(0.0052)	(0.0141)	(0.0191)	(0.0038)		(23.6109)	(0.1183)	(0.5261)	
SPF 4Q ahead recession probability	0.0459***	0.0187***	0.8991***	0.1145***	0.0567***	5.6150	3.2028	-1.1800***	-14296.88	13.11
	(0.0073)	(0.0050)	(0.0135)	(0.0181)	(0.0137)	(7.2397)	(2.4615)	(0.2518)	(0.0206)	
Industrial production	0.0463***	0.0189***	0.8987***	0.1127***	-0.0452***	1	4.5673***	-0.0878	-14317.10	6.01
	(0.0073)	(0.0052)	(0.0141)	(0.0184)	(0.0123)		(1.2065)	(0.0949)	(0.3213)	
Industrial production (2-sided filter)	0.0463***	0.0195***	0.8957***	0.1147***	-0.1025***	2.4348	5.9212***	0.1111	-14309.31	9.56
	(0.0073)	(0.0051)	(0.0138)	(0.0181)	(0.0281)	(1.0990)	(2.4321)	(0.1396)		
ADS index	0.0462***	0.0172***	0.8991***	0.1139***	-0.3944***	1	5.7082***	-0.2397**	-14310.19	9.73
	(0.0073)	(0.0054)	(0.0141)	(0.0183)	(0.0816)		(1.4493)	(0.0910)	(0.3720)	
Housing starts	0.0470***	0.0189***	0.8948***	0.1159***	-0.0173***	3.0722	4.8291*	-0.0954	-14299.24	16.59
	(0.0073)	(0.0053)	(0.0144)	(0.0185)	(0.0037)	(1.2979)	(2.1952)	(0.0899)	(0.0001)	
Housing starts (2-sided filter)	0.0469***	0.0193***	0.8930***	0.1169***	-0.0212***	4.6220	4.6151	-0.0786	-14297.60	17.68
	(0.0073)	(0.0053)	(0.0146)	(0.0185)	(0.0060)	(2.5848)	(2.3427)	(0.0942)		
Term spread	0.0470***	0.0188***	0.8930***	0.1161***	-0.2448***	3.8431	5.5547	0.2309**	-14304.11	12.46
	(0.0073)	(0.0053)	(0.0149)	(0.0192)	(0.0458)	(3.8808)	(4.4948)	(0.1162)	(0.0507)	
Principal component 1	0.0461***	0.0152***	0.8931***	0.1214***	0.2282***	1	3.2526***	-0.2047**	-14292.44	17.32
	(0.0073)	(0.0054)	(0.0143)	(0.0188)	(0.0318)		(0.7021)	(0.0816)	(0.5575)	
Principal component 2	0.0465***	0.0213***	0.8940***	0.1138***	0.3757***	4.7833	4.3873	-0.1913**	-14297.12	15.19
	(0.0073)	(0.0051)	(0.0139)	(0.0182)	(0.1136)	(3.5549)	(2.9713)	(0.0872)	(0.0000)	
Principal component 3	0.0452***	0.0148***	0.8950***	0.1208***	0.4137***	7.1145	7.7395	-0.2081***	-14287.94	18.10
	(0.0073)	(0.0052)	(0.0140)	(0.0187)	(0.0706)	(5.0254)	(4.6447)	(0.0817)	(0.0029)	
GJR-GARCH(1,1)	0.0469***	0.0221***	0.9025***	0.1068***	-	-	-	0.8446***	-14327.62	-
	(0.0073)	(0.0050)	(0.0137)	(0.0181)				(0.0872)		

*Bollerslev-Wooldridge QMLE robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. For the weight parameters I test $\omega_i = 1$. A "1" in the table indicates ω_1 is fixed to 1 before estimation. LLF is the value of the log-likelihood function and VR is the variance ratio from Section 2.3 multiplied by 100. The MIDAS polynomial: $\log \tau_t = m + \theta \sum_{k=1}^{K} \varphi_k(\omega_1, \omega_2) X_{t-k}$. All models are estimated with a restricted ($\omega_1 = 1$) and an unrestricted weighting scheme. The model reported in the table is chosen based on a likelihood ratio test between the restricted and unrestricted specifications. The related p-value is reported below the LLF.*

2.5.2 Combining macroeconomic and sentiment indicators

In order to examine the combined and relative information content of macroeconomic variables and survey-based sentiment indicators I include them simultaneously in the same MIDAS polynomial. Table 2.4 reports the results. The significance of their respective coefficients (θ_1 and θ_2) can be used to assess whether both are simultaneously useful for explaining volatility. The variance ratios reveal whether the variables are able to explain more of the total conditional variance together than on their own. As the term spread can also be interpreted as a sentiment indicator, I include specifications where the term spread is combined with the other macroeconomic data. As a robustness check, to control for the information in realised volatility, I include results with three explanatory variables in the MIDAS polynomial in Appendix 2.C: realised volatility, a macroeconomic variable and a sentiment indicator. Overall, the results in this section are robust to including realised volatility, which even makes the distinction between leading and lagging indicators sharper.

For each variable I keep the earlier choice of a restricted or unrestricted weighting scheme but re-estimate the weight parameter(s).[30] It is interesting to re-examine the lead-lag relationship discussed in detail in Conrad and Loch (2014) when several variables are included simultaneously in the MIDAS polynomial. In most cases the weight parameter(s) are similar regardless of the other variables included (compare Table 2.3 with Table 2.4). For example, the hump-shaped weight profile for housing starts is similar across models (see also Figure 2.1). The four quarters ahead recession probability and the term spread occasionally get weighting schemes with all weight on distant lags (ω_1 and ω_2 very large), implying the other included variable captures information in more recent lags. The same is true when the two-sided filter for industrial production is combined with the Buying conditions index, the ISM Recession indicator or the 1Q ahead recession probability. The fact that all weight for industrial production is on distant lags implies the leads are not behind the improvements in in-sample fit for these models. However, for several specifications significant weight is given to the leads (see Figure 2.1) and in fact when combined with the ISM New Orders index almost all weight is on the future expected values of industrial production. This also implies the ISM New Orders index captures the information in current and past industrial production. For the one quarter ahead recession probability all the weight is on the first lag ($\omega_2 > 100$ in Table 2.4 and Table 2.12). This is not surprising as the information in the lags of the one quarter ahead recession probability can plausibly be assumed to be already included in other data.

The intuitive sign of θ is retained in most cases where the parameter is statistically significant. The effect of industrial production[31] or the ADS index on long-term volatility is insignificant or only weakly significant when survey-based sentiment measures are included, with the exception of the four quarters ahead recession probability, while the sentiment indicators remain highly significant. This implies sentiment indicators capture information in and beyond backward-looking macroeconomic variables.

[30] The exception is the model combining information from housing starts and the term spread, where using an unrestricted weighting scheme for the term spread leads to the same model as when the restricted weighting scheme is used. Hence the table reports the model using the restricted weighting scheme.

[31] For results on the one-sided filter, see Appendix 2.E.

Table 2.4: Estimation results of the GARCH-MIDAS-X_1-X_2 model

Variables ($X_1 + X_2$)	θ_1	θ_2	ω_{11}	ω_{12}	ω_{21}	ω_{22}	m	LLF	VR
IP + ISM New Orders index	0.0551**	-0.0541***	6.3568	32.3254	1	3.2332*	3.0097***	-14297.83	15.25
	(0.0257)	(0.0145)	(7.3637)	(22.8047)		(1.1234)	(0.8687)		
IP + ISM Recession indicator	-0.0174**	-0.0555***	242.4155**	300**	1	2.8020*	0.3136**	-14301.99	12.10
	(0.0084)	(0.0121)	(116.3788)	(143.4724)		(0.9969)	(0.1252)		
IP + Consumer confidence	-0.0396	-0.1228***	3.6154	7.5122	1.5177	2.2555*	-0.0756	-14302.41	13.51
	(0.0453)	(0.0508)	(5.7807)	(7.8145)	(0.4816)	(0.7196)	(0.1816)		
IP + News Heard index	-0.0286	-0.0576***	4.1869	8.4485	1.6514*	1.7055**	-0.1114	-14300.19	14.71
	(0.0326)	(0.0202)	(4.8647)	(7.2544)	(0.3609)	(0.2996)	(0.1363)		
IP + Buying conditions index	-0.0132*	-0.1100***	237.5957***	300***	1	2.1761***	-0.1522**	-14290.36	18.61
	(0.0074)	(0.0196)	(87.5787)	(98.5534)		(0.3786)	(0.0787)		
IP + SPF 1Q ahead	-0.0205***	0.0142***	190.6479	240.8792	1	103.6299**	-0.4246***	-14292.83	15.23
	(0.0068)	(0.0020)	(141.2588)	(178.3991)		(50.7767)	(0.0910)		
IP + SPF 4Q ahead	-0.0679**	0.0538***	2.9914	5.1746	1.0000	1.0812	-0.9588***	-14285.45	16.73
	(0.0313)	(0.0189)	(2.2318)	(3.6924)	(2.5503)	(1.3931)	(0.3447)		
IP + Term spread	-0.1199***	-0.2752***	1.3288	6.2245	2.6769	1.9239	0.6708***	-14290.98	19.93
	(0.0352)	(0.0739)	(0.8646)	(3.2435)	(4.9670)	(3.9326)	(0.2347)		
ADS + ISM New Orders index	-0.0476	-0.0448***	1	112.4430***	1	4.3555*	2.2511**	-14302.50	13.16
	(0.0658)	(0.0101)		(19.7338)		(1.7870)	(0.5659)		
ADS + ISM Recession indicator	-0.2173*	-0.0461***	1	3.9159***	1	2.5149***	0.1670	-14303.78	11.90
	(0.1085)	(0.0148)		(0.9516)		(0.4976)	(0.1598)		
ADS + Consumer confidence	-0.2134*	-0.1190***	1	4.5938*	1.6389	2.7679*	-0.2124**	-14301.47	14.15
	(0.1252)	(0.0423)		(2.1422)	(0.4684)	(1.0520)	(0.0885)		
ADS + News Heard index	-0.1805	-0.0703***	1	5.8157	1.8336**	2.7253***	-0.2153**	-14296.71	17.20
	(0.1218)	(0.0181)		(5.5001)	(0.3502)	(0.6615)	(0.0859)		
ADS + Buying conditions index	-0.0978	-0.1084***	1	4.8705*	1	2.0713***	-0.1945**	-14292.44	17.92
	(0.1132)	(0.0273)		(2.0590)		(0.3123)	(0.0827)		
ADS + SPF 1Q ahead	-0.1949*	0.0121***	1	3.7040**	1	118.0634***	-0.4537***	-14297.70	13.76
	(0.1084)	(0.0029)		(1.0652)		(38.8150)	(0.0933)		
ADS + SPF 4Q ahead	-0.3083***	0.0388***	1	7.9303**	80.8063	20.9150	-0.9216***	-14286.06	17.41
	(0.0766)	(0.0095)		(2.5279)	(114.0771)	(26.8410)	(0.1773)		
ADS + Term spread	-0.2375	-0.2016***	1	6.2223**	1.3275	2.0194	0.1215	-14298.45	14.39
	(0.1613)	(0.0726)		(2.2110)	(5.3508)	(7.0048)	(0.1489)		

Table 2.4 *continued*

Variables ($X_1 + X_2$)	θ_1	θ_2	ω_{11}	ω_{12}	ω_{21}	ω_{22}	m	LLF	VR
Housing starts + ISM New Orders	-0.0119***	-0.0322*	2.6841	4.5192	1	3.6757	1.6275*	-14291.66	18.61
	(0.0038)	(0.0174)	(1.6222)	(3.7801)		(5.4495)	(0.9449)		
Housing starts + ISM Recession ind.	-0.0133***	-0.0422***	3.2868	4.3331	1	2.8435	0.2292*	-14289.97	19.32
	(0.0046)	(0.0132)	(2.1840)	(3.8762)		(1.9898)	(0.1250)		
Housing starts + Consumer confidence	-0.0157***	-0.0505	3.3762	4.4388	3.5668	14.0529	-0.1049	-14294.06	20.30
	(0.0050)	(0.0633)	(2.2100)	(3.7908)	(5.8869)	(37.1707)	(0.0896)		
Housing starts + News Heard index	-0.0183***	-0.0233	3.8926	6.6469	4.2756*	18.9465*	-0.0968	-14290.02	23.61
	(0.0036)	(0.0146)	(1.6349)	(3.5823)	(2.3486)	(16.6934)	(0.0856)		
Housing starts + Buying conditions	-0.0079**	-0.0862***	4.5095	7.3244	1	2.1633**	-0.1443*	-14287.77	21.43
	(0.0038)	(0.0242)	(2.2030)	(4.3348)		(0.5190)	(0.0810)		
Housing starts + SPF 1Q ahead	-0.0109***	0.0113***	4.8659	6.8390	1	130.8206***	-0.3613***	-14285.01	20.12
	(0.0032)	(0.0024)	(2.5747)	(3.8847)		(11.9711)	(0.1023)		
Housing starts + SPF 4Q ahead	-0.0115***	0.0391***	3.4720*	6.0574**	7.4149	3.6789	-0.8166***	-14284.51	18.24
	(0.0036)	(0.0141)	(1.4878)	(2.4506)	(14.1629)	(3.5301)	(0.2658)		
Housing starts + Term spread	-0.0124***	-0.2001***	3.3565*	5.8773*	1	1.0878	0.2119*	-14288.94	20.16
	(0.0032)	(0.0555)	(1.3815)	(2.6210)		(0.6190)	(0.1156)		
Term spread + ISM New Orders	-0.1653**	-0.0293	1.2356	1.7638	1	6.1276	1.6906	-14296.97	15.31
	(0.0770)	(0.0265)	(9.0325)	(11.1379)		(4.0488)	(1.4912)		
Term spread + ISM Recession ind.	-0.1358***	-0.0426**	166.57	121.8660	1	3.5158***	0.3956***	-14299.30	14.16
	(0.0486)	(0.0126)	(365.8079)	(257.6337)		(0.9715)	(0.1164)		
Term spread + Consumer confidence	-0.2132**	-0.0995**	2.7579	2.4139	2.2537	5.5324	0.1729	-14293.04	19.08
	(0.1033)	(0.0483)	(10.3230)	(9.6367)	(1.2068)	(5.1890)	(0.1960)		
Term spread + News Heard index	-0.2364***	-0.0692***	1.2576	1.1650	1.9056***	3.3570***	-0.2021	-14288.96	22.84
	(0.0688)	(0.0160)	(1.2969)	(1.1253)	(0.2802)	(0.7393)	(0.1357)		
Term spread + Buying conditions	-0.0894*	-0.0971***	198.1174	143.5106	1	2.3099***	-0.0332	-14289.28	19.56
	(0.0488)	(0.0231)	(148.2565)	(91.7209)		(0.5050)	(0.1047)		
Term spread + SPF 1Q ahead	-0.1293**	0.0108***	12.9660	13.2133	1	196.4147***	-0.1847	-14294.01	15.25
	(0.0532)	(0.0031)	(12.9660)	(15.1806)		(27.6688)	(0.1601)		
Term spread + SPF 4Q ahead	-0.1790***	0.0471***	8.1021	9.3810	1.1869	1.3657	-0.7171**	-14286.76	18.04
	(0.0589)	(0.0151)	(9.4493)	(9.0588)	(1.2928)	(0.7705)	(0.3180)		
PC 1 + PC 2	0.1726***	0.1132***	1	5.9045**	300***	184.2374***	-0.2062***	-14283.62	19.61
	(0.0368)	(0.0371)		(1.9298)	(101.4750)	(60.1322)	(0.0797)		
PC 1 + PC 3	0.1145**	0.2654***	1	3.5915	8.0493	8.5458	-0.2115***	-14282.10	19.55
	(0.0520)	(0.0983)		(2.0362)	(9.8807)	(8.1340)	(0.0783)		
PC 2 + PC 3	0.0870	0.3270	300**	183.9175***	13.1126	22.0828	-0.2092***	-14282.90	19.31
	(0.0563)	(0.2385)	(116.3457)	(67.3355)	(74.4119)	(107.0288)	(0.0804)		

*Bollerslev-Wooldridge QMLE robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. For the weight parameters I test $\omega_i = 1$. A "1" in the table indicates ω_i is fixed to 1 before estimation. LLF is the value of the log-likelihood function and VR is the variance ratio from Section 2.3 multiplied by 100. See notes on Table 2.1 for full names of explanatory variables. The MIDAS polynomial: $\log \tau_t = m + \theta_1 \sum_{k=1}^{K} \varphi_k(\omega_{11}, \omega_{12}) X_{1,t-k} + \theta_2 \sum_{k=1}^{K} \varphi_k(\omega_{21}, \omega_{22}) X_{2,t-k}$, where X_1 denotes the macroeconomic data and X_2 the sentiment data, as stated in the first column. For industrial production growth the two-sided filter with SPF forecasts is used.*

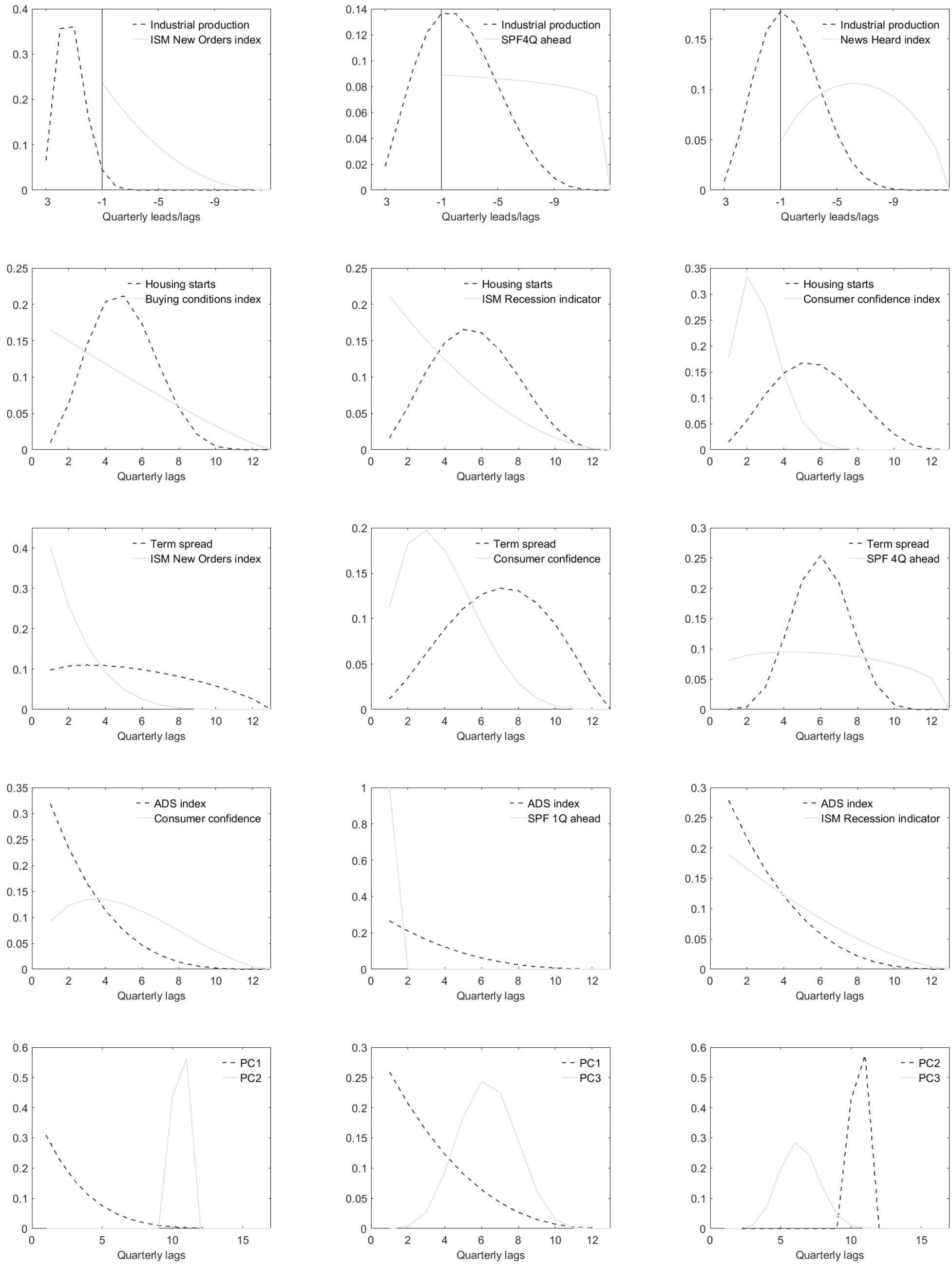


Figure 2.1: MIDAS weight profiles of selected models from Table 2.4.

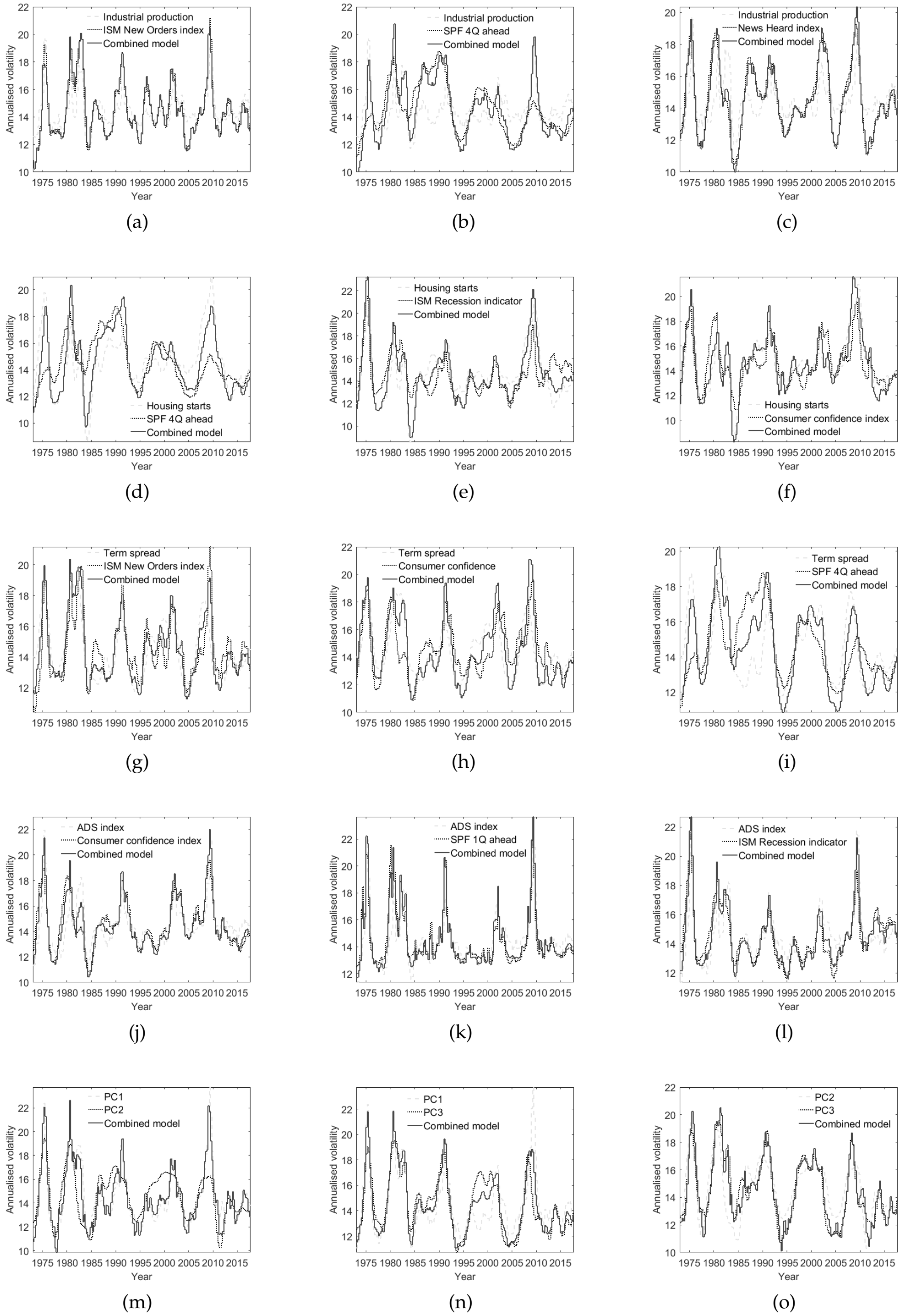


Figure 2.2: Long-term volatility components of selected GARCH-MIDAS-X_1-X_2 models (from Table 2.4) ('Combined model') and corresponding GARCH-MIDAS-X models (Table 2.3).

On the other hand, when the two-sided filter is used for industrial production the importance of the macroeconomic variable strengthens. This is in particular true when combined with the term spread, as now industrial production gets a negative and statistically significant parameter estimate, and the in-sample fit clearly improves.

Housing starts gets highly significant estimates for θ, and with the exception of some consumer confidence indicators, so does the sentiment data. The term spread and the sentiment indicators are simultaneously highly significant, but when information in the absolute value of returns is taken into account there is much weaker evidence of sentiment indicators containing additional useful information for long-term stock market volatility (Table 2.12). It is noteworthy that the four quarters ahead recession probability is highly significant throughout the specifications, indicating it includes information complementary to that in macroeconomic data (and realised volatility).

From Figure 2.2 we can see that when only one of the variables is significant in the GARCH-MIDAS-X_1-X_2 model the long-term volatility component follows closely the long-term volatility component of this individual model (for example, Figure 2.2a). When both variables get significant parameter estimates also the long-term volatility component is a combination of the long-term volatility components of the individual models (for example, Figure 2.2i).

Comparing the variance ratios in Table 2.4 to those in Table 2.3 we can see that especially when only one of the variables is significant the gains from including both macroeconomic and sentiment data in the same model can be small or non-existent. On the other hand, when both variables are simultaneously significant there are naturally clearer gains: the variance ratio for the model based on the term spread is 12.5%, while the variance ratio for the model with the consumer confidence index is 12.1% (Table 2.3). The long-term volatility component based on a combination of these series explains over 19% of the total variance (Table 2.4). The model driven by only industrial production has a variance ratio of roughly 9%, but when combined with the term spread the variance ratio rises to almost 20%. The principal components summarise information in the economy well leading to some of the highest variance ratios of close to 20% (Table 2.4).

2.6 Out-of-sample results

The in-sample results highlighted the superiority of the absolute value of returns and forward-looking data for stock return volatility. In this section I consider whether these results extend to an out-of-sample context. I start by determining whether the GARCH-MIDAS-X models outperform the GJR-GARCH(1,1) model in forecasting (total) volatility. In order to narrow down the set of potential forecasting models for stock return volatility I apply the Model Confidence Set (MCS) procedure by Hansen et al. (2011) to the GARCH-MIDAS-X models. In Section 2.6.2 I examine whether the models combining information from macroeconomic and sentiment data are superior in an out-of-sample context to models with only one explanatory variable.

As the short-term GARCH components are similar across all GARCH-MIDAS specifications, the largest gains in forecasting can be achieved in long-term forecasts. Thus, I consider forecasts for 1, 2, 3 and 4 quarters ahead. The one-step ahead volatility prediction is given directly by equations (2.1) and (2.2). For further horizons we need to iterate forward the daily GJR-GARCH model forecasts and combine this short-term forecast with a forecast for the long-term component, τ_t. For the GJR-GARCH model the forecast for day i is formed as:

$$E\left[g_{i,t}|F_{N_{t-1},t-1}\right] = 1 + (\alpha + \beta + \gamma/2)^{i-1}(g_{1,t} - 1), \tag{2.4}$$

where N_t is the number of trading days in period t and $F_{N_{t-1},t-1}$ is the information set in period $t-1$. The forecast for total volatility for period t can be expressed as:

$$E\left[\sum_{i=1}^{N_t} g_{i,t}\tau_t\varepsilon_{i,t}^2|F_{N_{t-1},t-1}\right] = \tau_t\left[N_t + (g_{1,t} - 1)\frac{1-(\alpha+\beta+\gamma/2)^{N_t}}{1-\alpha-\beta-\gamma/2}\right]. \tag{2.5}$$

This forecast can be iterated for any forecast horizon. Following Conrad and Loch (2014) I create non-overlapping quarterly forecasts by summing the daily forecasts over the respective quarter while keeping τ_t fixed at its one-step ahead prediction for all horizons.[32] Because the forecast of the GARCH component converges to its (constant) unconditional expectation as the forecast horizon increases, in the long run the forecast differences of the different models are entirely driven by the long-term components.

The first estimation period is Q1 1973 - Q4 1998, and the out-of-sample period is Q1 2000 - Q3 2017 (71 quarters). The forecasts are evaluated against realised volatility calculated as the sum of squared daily returns ($RV_t = \sum_{i=1}^{N_t} r_{i,t}^2$). I use a rolling window estimation scheme moving forward the estimation period one quarter every period. For each variable I keep the earlier choice of a restricted or unrestricted weighting scheme but re-estimate the weight parameter(s) each period. The principal components are calculated recursively, including data only up until the end of each estimation period. I compare the forecast accuracy of the models using the ratio between their mean absolute forecast errors (MAFE) or their mean squared forecast errors (MSFE). The MSFEs penalise relatively more heavily the largest forecast errors, which is useful if one wants to emphasise large individual forecast errors over smaller ones. The forecast errors during the financial crisis are in particular emphasised by the MSFE, and as a robustness check I include results excluding the large forecast errors occuring in conjunction with the financial crisis in Appendix 2.F. Poon and Granger (2003) noted that when using squared returns as the quantity of interest and using squared errors as the measure of forecast accuracy one is effectively comparing the fourth moments of the data, which can complicate the comparison. However, Patton (2011) argued that while the MSFE loss function is robust in the sense that using a noisy proxy for volatility (such as the sum of squared daily returns) does not change the ranking of forecasting models, the MAFE loss function is not. In general, the ranking of the models is similar but statistical significance is weaker when using squared forecast errors. The statistical significance of the differences in forecasting performance is assessed using the (unconditional) predictive ability test by Giacomini and White (2006).

The Model Confidence Set (MCS) procedure by Hansen et al. (2011) determines

[32] The long-term component could be forecast using forecasts for the explanatory data. However, using the SPF forecasts for industrial production and housing starts as forecasts in equation (2.5) does not significantly alter the volatility predictions. These results are available upon request.

the set of models that includes the best model(s) at a pre-specified level of confidence out of an initial set of candidate models.[33] It tests whether all models are equally good (equivalence test) based on their relative performance, and if not, the model with the worst sample performance is eliminated (based on an elimination rule). Hence it does not require choosing a benchmark model. The process is repeated until equal performance cannot be rejected among the remaining models at the pre-specified level of confidence. The surviving models form the MCS which includes the best model(s) with a certain probability.

2.6.1 Forecasting with the GARCH-MIDAS-X model

It is clear from Table 2.5 that the GJR-GARCH(1,1) model is significantly more difficult to outperform than the GARCH-MIDAS-RV model. One quarter ahead none of the GARCH-MIDAS models perform significantly better than the benchmark, with the GARCH-MIDAS model driven by the first principal component being the best model.[34] The one quarter ahead recession probability also performs well for the shortest horizon, indicating that knowledge of whether we are heading for a recession in the near future is useful for forecasting volatility. At longer horizons the term spread, the consumer confidence subindices, housing starts as well as the second and third PC outperform the GJR-GARCH(1,1) model, but the differences are only statistically significant when we look at the absolute value of the forecast errors. The realised volatility measure based on the sum of the absolute value of returns performs better than squared returns also out-of-sample while the four quarters ahead recession probability performs relatively poorly compared to the baseline GJR-GARCH model. Most GARCH-MIDAS specifications produce lower average forecast errors than the benchmark at some horizons at least, although the differences tend not to be statistically significant. This is especially true for the MSFE ratios.

The MCS procedure can only separate between models if the data (here the mean absolute or squared forecast errors) is informative enough. The longer the forecasting horizon the more important is the long-term component, and the more the forecast errors vary between models. This is reflected in Table 2.5, where at the one and two quarters ahead horizons all models are chosen into the 90% confidence set (bold numbers), while at the three and four quarters ahead horizons the GARCH-MIDAS models driven by the term spread, the realised volatility measures, the four quarters ahead recession probability and the third principal component (which is mostly correlated with the term spread and the four quarters ahead recession probability) are included in the set of superior models when considering absolute forecast errors. The MCS thus favours forward-looking variables, and in particular the term spread, as predictors for stock return volatility over long horizons. The inclusion of the realised volatility measures is surprising considering their weak forecasting performance compared to

[33]For technical details on the MCS procedure and its bootstrap implementation, see Appendix 2.G and Hansen et al. (2011). The MCS procedure is implemented using the R package MCS, see Bernardi and Catania (2018), with $B = 10000$ and block bootstrap. Confidence level: $\alpha = 0.10$.

[34]This conclusion partly contrasts the one in Asgharian et al. (2013), where the GARCH(1,1) model was the worst model one period ahead when models were ranked by total volatility, although the differences were not statistically significant.

Table 2.5: Forecasting performance of the GARCH-MIDAS-X model

	1 quarter ahead		2 quarters ahead		3 quarters ahead		4 quarters ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Realised volatility (Σr^2)	**1.21**	**1.63**	**1.28**	**1.49**	**1.27**	**1.53**	**1.26**	**1.51**
Realised volatility ($\Sigma\|r\|$)	**1.13**	**1.38***	**1.22**	**1.38**	**1.24**	**1.42**	**1.24**	**1.43**
Consumer confidence	**0.98**	**0.95**	**0.97**	**0.94**	0.95	**0.95**	0.96	**0.97**
News Heard index	**0.98**	**1.08**	**0.95**	**0.96**	0.92**	**0.95**	0.91**	**0.95**
Buying conditions index	**0.97**	**1.17**	**0.93****	**0.89**	0.92**	**0.91**	0.93*	**0.93**
ISM New Orders index	**0.99**	**1.02**	**0.99**	**1.01**	0.97	**0.99**	0.96*	**0.99**
ISM Recession indicator	**1.05**	**0.98**	**1.07**	**0.99**	1.07	**0.99**	1.04	**1.00**
SPF 1Q ahead	**0.96**	**0.87**	**0.99**	**0.99**	1.00	**0.98**	0.99	**1.00**
SPF 4Q ahead	**1.02**	**1.14***	**0.99**	**1.06***	0.96	**1.04**	**0.94**	**1.03**
IP (2-sided filter)	**1.01**	**0.94**	**1.03**	**0.99**	1.02	**0.99**	0.98	**0.98**
ADS index	**0.99**	**0.84**	**1.01**	**0.97**	0.99	**0.98**	0.98	**0.98**
Housing starts	**0.99**	**0.98**	**0.95**	**0.93**	0.93*	**0.92**	0.93**	**0.92**
Term spread	**1.05**	**1.44**	**0.93****	**0.98**	**0.87*****	**0.92**	**0.85*****	**0.91**
Principal component 1	**0.94**	**0.77**	**0.96**	**0.92**	0.96	**0.94**	0.98	**0.96**
Principal component 2	**1.02**	**1.19**	0.96	**1.01**	0.93***	**0.98**	0.92***	**0.97**
Principal component 3	**1.09**	**1.46**	**0.96**	**0.98**	**0.89*****	**0.92**	**0.88*****	**0.92**

*Benchmark model: GJR-GARCH(1,1). MAFE ratio: $\frac{MAFE_{GMX}}{MAFE_{GARCH}}$, where GMX stands for the GARCH-MIDAS-X model. The MSFE ratio is calculated equivalently. A value below 1 means the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5% and 1% level, respectively, according to the Giacomini and White (2006) test. A bolded value indicates the forecast is included in the model confidence set ($\hat{M}^*_{90\%}$) with $\alpha = 0.10$.*

the GJR-GARCH model. When looking at squared forecast errors the MCS cannot separate between the models at any horizon.

2.6.2 Forecasting with the GARCH-MIDAS-X_1-X_2 model

To determine whether out-of-sample forecasting performance can be improved by combining information in macroeconomic and sentiment data I compare the forecasting performance of the GARCH-MIDAS-X_1-X_2 models (from Section 2.5.2) to the more parsimonious GARCH-MIDAS-X models (from Section 2.5.1) using both squared (Table 2.6) and absolute (Table 2.7) forecast errors.[35] The benchmark model in the top panel is the GARCH-MIDAS model driven by macroeconomic data, and the panel thus describes the marginal benefit of using a GARCH-MIDAS model combining macroeconomic and sentiment data over a model only driven by macroeconomic data. The benchmark model in the middle panel is the GARCH-MIDAS model driven by sentiment data, and the panel thus describes the additional explanatory power arising from adding macroeconomic data to a GARCH-MIDAS model driven by sentiment data. The bottom panel considers models combining information in principal components.

[35]Forecast accuracy could also be directly compared to the GJR-GARCH model, but this would give us a less clear picture of the relative importance of the different types of variables.

Including the term spread, housing starts or industrial production (two-sided filter)[36] in the long-term component is beneficial from a forecasting perspective, as reflected in MSFE ratios of less than one (middle panel), as well as several statistically significant improvements when the absolute value of the forecast errors is used. The term spread does, however, perform poorly on the one quarter ahead forecasting horizon, while the ADS index and industrial production seem mostly useful on this horizon. On the other hand, the term spread, consumer confidence indicators and the one quarter ahead recession probability (one the one quarter ahead horizon) tend to improve forecasts compared to models only driven by macroeconomic data (top panel), but for the MSFEs these differences are mostly not statistically significant. When using the absolute value loss function statistical significance is stronger. The four quarters ahead recession probability performs surprisingly poorly, as adding it to a model tends to lead to weaker forecasting performance. This indicates that anticipating an oncoming recession far in the future cannot be used to improve volatility forecasts compared to models driven by other data. The ISM indices are only rarely useful for forecasting when other data is available. For the principal components the main takeaway is that both PC1 and PC3 include complementary information to PC2, while adding PC1 is especially useful over short horizons.

If we compare the MSFE ratios of the top and middle panels of Table 2.6, we see that in several cases, most prevalently for industrial production, housing starts and the term spread, the corresponding ratios in the two panels are both less than one (for example, first entry in top panel (0.84) versus first entry in middle panel (0.83)). This indicates that the model combining both macroeconomic and survey-based sentiment data improve on both the underlying simpler models, making the GARCH-MIDAS-X_1-X_2 model superior to both nested GARCH-MIDAS-X models. However, both improvements are never statistically significant simultaneously.

[36]The benefit of adding industrial production to a model disappears when the one-sided filter is used, see Appendix 2.E. The difference between the one and two-sided filters when adding sentiment data to the model driven by industrial production (top panel) is less clear as the same macroeconomic variable is now included in the models in both the numerator and the denominator.

Table 2.6: MSFE ratios: GARCH-MIDAS-X_1-X_2 model vs. GARCH-MIDAS-X model

Benchmark: GARCH-MIDAS-X model, where X is macroeconomic data (as indicated in the first row)

	Industrial production (2-sided)				ADS index				Housing starts				Term spread			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
CC	0.84	0.95*	0.98	0.99	1.08	0.98	0.97	0.99	0.87	0.98	0.98	1.00	0.93	0.96	0.97	0.99
NH	1.24	1.01	0.98	0.98*	1.26	0.99	0.97	0.98	1.06	1.02	1.01	0.99	0.97*	0.95	0.97	0.99
BC	1.14	0.88	0.91	0.94	1.42	0.92	0.93	0.96	1.17	0.95	0.98	1.01	1.04	0.92	0.95	0.97
ISM NO	0.97	1.00	1.00	0.99	1.14	1.03*	1.03*	1.02	1.10	1.04*	1.03	1.03	0.96	1.05	1.06	1.05
ISM RI	0.96	0.98	0.98	0.98	1.35	1.05*	1.02	1.02	1.01	1.04***	1.04*	1.04	0.92	1.01	1.00	1.02
SPF 1Q	0.78	0.99	0.97	1.00	1.01	1.01	0.99	1.01	0.86	1.03*	1.01	1.03	0.88	0.99	0.99	1.01
SPF 4Q	1.17	1.05	1.04	1.02	1.21	1.07*	1.05	1.05	1.22	1.13	1.11	1.10	1.04***	1.06	1.07	1.07
TS	1.17	0.91	0.88	0.90	1.59	0.99	0.94	0.93	1.44	1.02	0.98	0.97				

Benchmark: GARCH-MIDAS-X model, where X is sentiment data (as indicated in the first column)

	Industrial production (2-sided)				ADS index				Housing starts				Term spread			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
CC	0.83	1.00	1.02*	1.01	0.95	1.01	1.00	1.00	0.90	0.97	0.94	0.94	1.41	1.00	0.94	0.93
NH	1.04	1.02	1.01	1.01	0.98	1.00	1.00	1.01	0.98	0.99	0.98	0.95	1.29	0.96	0.94	0.95
BC	0.92	0.98	0.99	0.99	1.02	1.01	1.01	1.01	0.98	0.99	0.99*	1.00	1.28	1.01	0.96	0.95
ISM NO	0.89	0.99	0.99	0.98	0.94	1.00	1.01	1.01	1.06	0.96	0.95	0.95	1.36	1.02	0.98***	0.96**
ISM RI	0.92	0.99	0.98	0.97	1.15	1.03	1.01	1.01	1.01	0.97	0.96	0.96	1.34	1.00	0.93***	0.93**
SPF 1Q	0.85	0.99	0.98	0.98	0.97	0.99	0.99	0.99	0.97	0.96	0.95	0.95	1.46	0.97	0.94	0.92
SPF 4Q	0.97	0.99	0.99	0.97*	0.89	0.98	0.99	0.99	1.05	0.98	0.97	0.97	1.32	0.98	0.95	0.95
TS	0.76	0.93	0.95	0.97	0.92	0.99	1.00	1.00	0.98**	0.97	0.97	0.98				

Benchmark: GARCH-MIDAS-X model, where X is a principal component (as indicated in the first column)

	Principal component 1				Principal component 2				Principal component 3			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
PC 1					1.22	1.03	0.99	0.99	1.87	1.12	1.01	1.02
PC 2	0.79	0.93	0.96	0.98					1.35	0.94	0.93	0.95*
PC 3	0.98	1.04	1.03	1.06	1.10	0.96	0.99	1.00				

*Top panel: MSFE ratio of a GARCH-MIDAS-X_1-X_2 model including macroeconomic and sentiment data and a GARCH-MIDAS-X model including macroeconomic data: $\frac{MSFE_{macro+sentiment}}{MSFE_{macro}}$. Middle panel: MSFE ratio of a GARCH-MIDAS-X_1-X_2 model including macroeconomic and sentiment data and a GARCH-MIDAS-X model including sentiment data: $\frac{MSFE_{macro+sentiment}}{MSFE_{sentiment}}$. A value below 1 means the model combining macroeconomic and sentiment data outperforms the model driven by macroeconomic or sentiment data. Bottom panel: MSFE ratio of a GARCH-MIDAS-X_1-X_2 model including two principal components and a GARCH-MIDAS-X model, where X is the principal component corresponding to the first column: $\frac{MSFE_{PC_{row}+PC_{column}}}{MSFE_{PC_{row}}}$. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5% and 1% level, respectively, according to the Giacomini and White (2006) test. See notes on Table 2.1 for full names of explanatory variables.*

Table 2.7: MAFE ratios: GARCH-MIDAS-X_1-X_2 model vs. GARCH-MIDAS-X model

Benchmark: GARCH-MIDAS-X model, where X is macroeconomic data (as indicated in the first row)

	Industrial production (2-sided)				ADS index				Housing starts				Term spread			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
CC	0.94*	0.95*	0.95*	0.98	0.99	0.96	0.96*	0.97	0.99	0.98	0.99	1.01	0.99	0.99	1.00	1.02
NH	1.03	0.96	0.93**	0.94***	1.00	0.94*	0.92***	0.93**	1.01	1.00	0.96	0.97	0.98	0.98	1.00	1.02
BC	0.96	0.92**	0.91***	0.95*	1.00	0.94*	0.92**	0.95	0.99	0.98	0.97	0.99	0.99	0.98	0.99	1.02
ISM NO	0.96*	0.96**	0.96*	0.96*	0.99	0.99	0.98	0.98	1.02	1.02	0.99	0.98	1.00	1.05***	1.05**	1.05**
ISM RI	1.02	1.05	1.05	1.05	1.08	1.06	1.06	1.06	1.02	1.04	1.06	1.04	1.00	1.04*	1.05*	1.07***
SPF 1Q	0.93	0.96	0.98	1.01	0.95	0.97	0.99	1.00	0.95	1.00	1.03	1.02	0.97	1.00	1.03	1.04*
SPF 4Q	1.01	0.97	0.95	0.94	1.02	0.98	0.98	0.97	1.03	1.05	1.01	0.99	1.04*	1.05**	1.05**	1.05
TS	0.97	0.88***	0.86***	0.87***	1.04	0.93**	0.88***	0.87***	1.03	0.97	0.92***	0.91***				

Benchmark: GARCH-MIDAS-X model, where X is sentiment data (as indicated in the first column)

	Industrial production (2-sided)				ADS index				Housing starts				Term spread			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
CC	0.97	1.00	1.02	1.01	0.97	1.00	1.02	1.01	1.01	0.97	0.97	0.97	1.07	0.94**	0.91***	0.90***
NH	1.04**	1.03**	1.02	1.00	1.02	1.00	0.99	1.00	1.02	0.99	0.97	0.97	1.05	0.95**	0.94***	0.94**
BC	1.01	1.01	1.01	1.00	1.03*	1.01	1.00	1.00	1.02	1.00	0.99	0.98	1.08*	0.97	0.94***	0.93***
ISM NO	0.98	0.99	1.00	0.98	1.00	1.00	1.00	1.00	1.03	0.98	0.95	0.95*	1.06	0.98	0.93***	0.92***
ISM RI	0.98	1.01	1.01	0.99	1.03	0.99	0.98	1.00	0.97	0.93**	0.92**	0.92*	1.01	0.90***	0.86***	0.87***
SPF 1Q	0.98	1.00	1.00	1.00	0.99	0.99	0.98	0.99	0.99	0.97	0.96*	0.95**	1.06	0.94**	0.89***	0.89***
SPF 4Q	1.01	1.00	1.01	0.98	0.99	1.00	1.01	1.01	1.01	1.00	0.98	0.97*	1.07	0.98	0.95**	0.94**
TS	0.93	0.97	1.01	1.01	0.98	1.00	1.01	1.01	0.98	1.00	0.99	1.00				

Benchmark: GARCH-MIDAS model, where X is a principal component (as indicated in the first column)

	Principal component 1				Principal component 2				Principal component 3			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
PC 1					1.04	0.97	0.97	0.96	1.16	1.02	0.96	0.92
PC 2	0.96	0.96	1.00	1.02					1.07	0.99	0.96	0.93**
PC 3	1.00	1.01	1.03	1.03	1.00	0.99	1.00	0.98				

*Top panel: MAFE ratio of a GARCH-MIDAS-X_1-X_2 model including macroeconomic and sentiment data and a GARCH-MIDAS-X model including macroeconomic data: $\frac{MAFE_{macro+sentiment}}{MAFE_{macro}}$. Middle panel: MAFE ratio of a GARCH-MIDAS-X_1-X_2 model including macroeconomic and sentiment data and a GARCH-MIDAS-X model including sentiment data: $\frac{MAFE_{macro+sentiment}}{MAFE_{sentiment}}$. A value below 1 means the model combining macroeconomic and sentiment data outperforms the model driven by macroeconomic or sentiment data. Bottom panel: MAFE ratio of a GARCH-MIDAS-X_1-X_2 model including two principal components and a GARCH-MIDAS-X model, where X is the principal component corresponding to the first column: $\frac{MAFE_{PC_{row}+PC_{column}}}{MAFE_{PC_{row}}}$. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5% and 1% level, respectively, according to the Giacomini and White (2006) test. See notes on Table 2.1 for full names of explanatory variables.*

2.7 Conclusion

This paper studies the relative and combined information content and predictive ability of macroeconomic variables and survey-based sentiment indicators. The paper explores whether sentiment indicators summarise the state of the economy and describe expectations of future macroeconomic conditions in a way that is either overlapping or complementary to macroeconomic fundamentals. In addition, I examine the information content in the absolute value of returns, forward-looking subcomponents of the consumer confidence index, and SPF recession probabilities for volatility. I also use a data set with real-time revisions to the lags of the macroeconomic variables taken into account in each period to match the information set available in real-time as closely as possible.

First of all, my results highlight the importance of the forward-looking subcomponents of the consumer confidence index. Over short-horizons the one quarter ahead recession probability performs well, while the four quarters ahead recession probability performs surprisingly poorly in out-of-sample forecasting. Secondly, once information in sentiment indicators is taken into account, backward-looking macroeconomic variables (for example, ADS index) contain only little additional information for modelling or forecasting volatility, while forward-looking variables (for example, term spread) remain useful, in particular over long horizons. Thirdly, there are several cases in which the GARCH-MIDAS-X_1-X_2 models lead to improvements in both in-sample fit and out-of-sample forecast accuracy compared to both nested GARCH-MIDAS-X models, although the differences in forecasting performance are not statistically significant.

Overall, the results support hte use of forward-looking variables for forecasting stock market volatility, but does not provide strong support for combining different types of data in one model. It is, however, also clear from the results that the standard asymmetric GARCH model is difficult to outperform in a statistically significant way, even over long horizons, by using economic data in a GARCH-MIDAS model.

References

Amendola, A., V. Candila, and G. M. Gallo (2019). On the asymmetric impact of macro–variables on volatility. *Economic Modelling 76*, 135–152.

Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold (2006). Volatility and correlation forecasting. *Handbook of Economic Forecasting 1*, 777–878.

Andreou, E., E. Ghysels, and A. Kourtellos (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics 158*, 246–261.

Arnold, I. J. and E. B. Vrugt (2008). Fundamental uncertainty and stock market volatility. *Applied Financial Economics 18*, 1425–1440.

Asgharian, H., C. Christiansen, and A. J. Hou (2015). Effects of macroeconomic uncertainty on the stock and bond markets. *Finance Research Letters 13*, 10–16.

Asgharian, H., A. J. Hou, and F. Javed (2013). The importance of the macroeconomic variables in forecasting stock return variance: a GARCH-MIDAS approach. *Journal of Forecasting 32*, 600–612.

Awartani, B. M. A. and V. Corradi (2005). Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries. *International Journal of Forecasting 21*, 167–183.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Bansal, R. and A. Yaron (2004). Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance 59*, 1481–1509.

Barsky, R. B. and E. R. Sims (2012). Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review 102*(4), 1343–1377.

Bauer, M. D. and T. M. Mertens (2018). Economic forecasts with the yield curve. *FRBSF Economic Letter 2018-07*.

Bernardi, M. and L. Catania (2018). The model confidence set package for R. *International Journal of Computational Economics and Econometrics 8*(2), 144–158.

Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the Business and Economic Statistics Section, Americal Statistical Association*, 177–181.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal 101*, 157–179.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies 1*, 195–228.

Campbell, S. D. and F. X. Diebold (2009). Stock returns and expected business conditions: Half a century of direct evidence. *Journal of Business & Economic Statistics 27*(2), 266–278.

Christiansen, C., M. Schmeling, and A. Schrimpf (2012). A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics 27*(6), 956–977.

Conrad, C. and O. Kleen (2019). Two are better than one: Volatility forecasting using multiplicative component GARCH-MIDAS models. *Journal of Applied Econometrics (forthcoming)*.

Conrad, C. and K. Loch (2014). Anticipating long-term stock market volatility. *Journal of Applied Econometrics 30*(7), 1090–1114.

Conrad, C. and M. Schienle (2018). Testing for an omitted multiplicative long-term component in GARCH models. *Journal of Business & Economic Statistics*, 1–14.

Davis, N. and A. M. Kutan (2003). Inflation and output as predictors of stock returns and volatility: international evidence. *Applied Financial Economics 13*(9), 693–700.

Diebold, F. X. and K. Yilmaz (2008). Macroeconomic volatility and stock market volatility, world-wide. *NBER Working Paper* (14269).

Ding, Z., C. W. Granger, and R. F. Engle (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance 1*, 83–106.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *The Review of Economics and Statistics 95*(3), 776–797.

Engle, R. F. and J. G. Rangel (2008). The Spline-GARCH model for low frequency volatility and its global macroeconomic causes. *Review of Financial Studies 21*, 1187–1222.

Errunza, V. and K. Hogan (1998). Macroeconomic determinants of European stock market volatility. *European Financial Management 4*(3), 361–377.

Fama, E. F. and K. R. French (1989). Business conditions and expected returns on stock and bonds. *Journal of Financial Economics 25*, 23–49.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The MIDAS touch: Mixed data sampling regression models. *EFA 2009 Bergen Meetings Paper*.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2005). There is a risk-return tradeoff after all. *Journal of Financial Economics* (76), 509–548.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics 131*(1-2), 59–95.

Ghysels, E., A. Sinko, and R. Valkanov (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*, 1545–1578.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Hamilton, J. D. and G. Lin (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics 11*(5), 573–593.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Mele, A. (2007). Asymmetric stock market volatility and the cyclical behavior of expected returns. *Journal of Financial Economics 86*, 446–478.

Officer, R. (1973). The variability of the market factor of the New York Stock Exchange. *Journal of Business 46*, 434–453.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics 160*, 246–256.

Paye, B. S. (2012). Deja vol: Predictive regressions for aggregate stock market volatility using macroeocnomic variables. *Journal of Financial Economics 106*, 527–546.

Pierdzioch, C., J. Dopke, and D. Hartmann (2008). Forecasting stock market volatility with macroeconomic variables in real time. *Journal of Economics and Business 60*, 256–276.

Poon, S.-H. and C. W. Granger (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature 41*, 478–539.

Romer, C. (1990). The great crash and the onset of the great depression. *The Quarterly Journal of Economics 105*, 597–624.

Schwert, W. G. (1989a). Business cycles, financial crises, and stock volatility. *Carnegie-Rochester Conference Series on Public Policy 31*, 83–126.

Schwert, W. G. (1989b). Why does stock market volatility change over time? *Journal of Finance 44*, 1115–1153.

Taylor, S. J. (1986). *Modelling financial time series*. New York: John Wiley & Sons.

Veronesi, P. (1999). Stock market overreaction to bad news in good times: a rational expectations equilibrium model. *Review of Financial Studies 12*, 975–1007.

Wang, F. and E. Ghysels (2015). Econometric analysis of volatility component models. *Econometric Theory 31*, 362–393.

Appendices

2.A Data description

Table 2.8: Descriptive statistics: Q1 1970 - Q3 2017

Variable	N	Mean	Std. dev.	Minimum	Maximum
CRSP daily returns	12 047	0.04	1.03	-17.41	11.35
Sum of squared returns	191	66.59	96.31	9.84	795.83
Sum of absolute value of returns	191	44.08	22.09	15.12	160.52
Consumer confidence index	191	0.08	5.19	-14.70	16.50
News Heard index	191	0.20	17.51	-59.00	52.00
Buying conditions index	191	0.28	8.42	-39.00	25.00
ISM New Orders index	191	54.94	7.46	27.27	71.90
ISM Recession indicator	191	8.34	6.64	-12.73	30.73
SPF 1Q ahead recession probability	191	18.83	16.05	2.16	74.78
SPF 4Q ahead recession probability	191	17.28	5.72	4.51	33.34
Industrial production (latest data)	191	2.22	6.04	-23.89	18.43
ADS index (latest data)	191	-0.10	0.81	-3.36	1.70
Housing starts (latest data)	191	5.34	36.77	-65.06	199.65
Term spread	191	1.73	1.21	-1.40	3.80

N denotes the Number of observations and Std. dev. the standard deviation.

Data sources:

- CRSP returns: Kenneth French's Data Library
- Realised volatility: calculated from the CRSP returns
- Consumer confidence data: University of Michigan (`http://www.sca.isr.umich.edu/`)
- ISM indices: FRED database (St. Louis Fed) and Institute for Supply Management (`https://www.instituteforsupplymanagement.org/`)
- Survey of Professional Forecasters data: Philadelphia Fed
- Industrial production, housing starts: Philadelphia Fed
- ADS index: Philadelphia Fed real-time center
- Term spread: FRED database (St. Louis Fed)

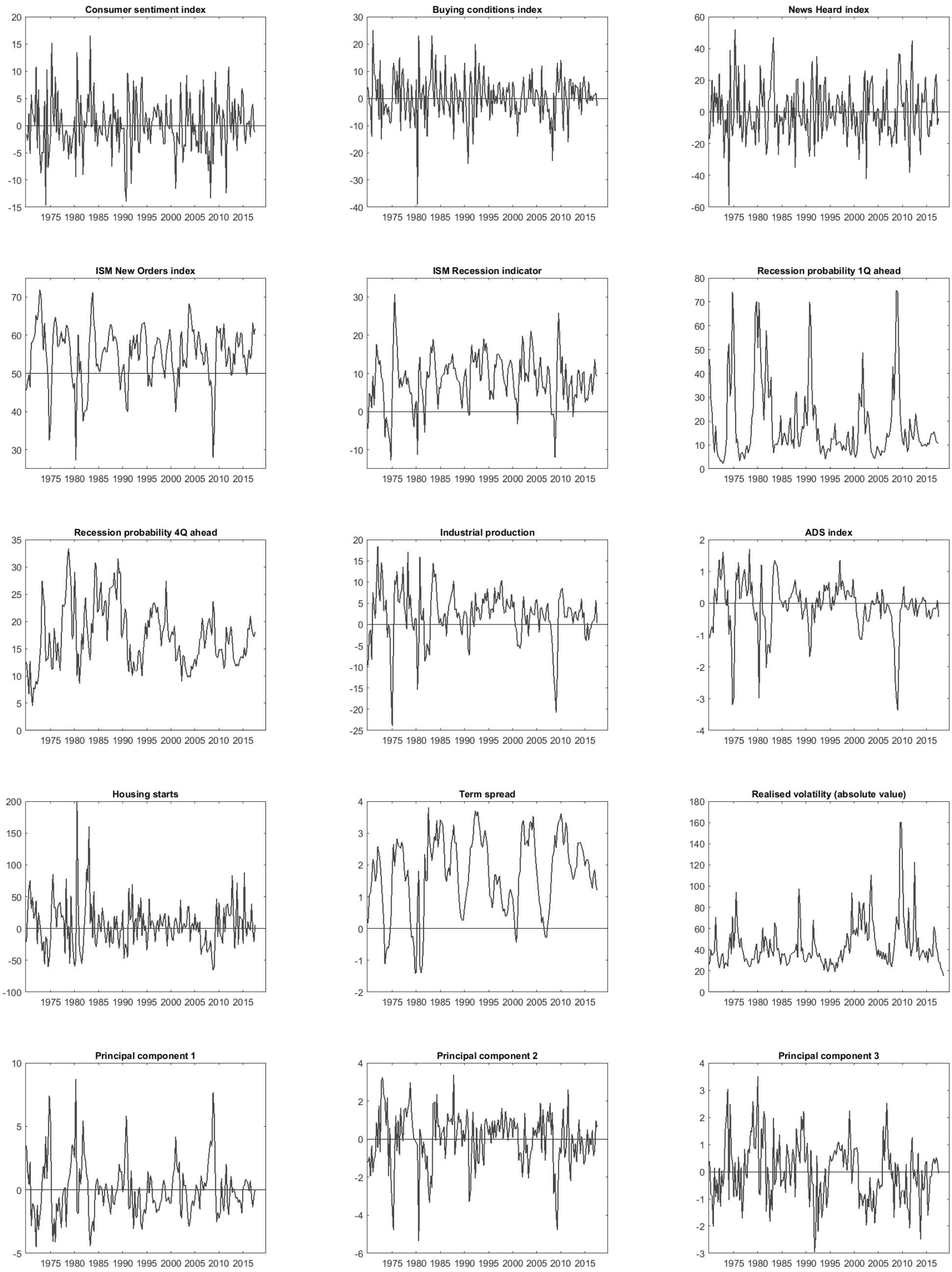


Figure 2.3: Explanatory data.

2.B Impact of revising lags in real time

I use a real-time data set for industrial production and housing starts where actual revisions are taken into account in the lags of the data. This appendix compares the in-sample fit and out-of-sample forecasts using first release data and data where the lags are revised.

The difference between using only first release data and using revised lags in the MIDAS polynomial is overall relatively small. As can be seen in Table 2.9, for industrial production growth the overall in-sample fit (measured by the variance ratio) is somewhat better when using first release data, while for housing starts the opposite is true. Looking at the variance ratio calculated over rolling windows for the forecasting sample (Figure 2.4), revised lags leads to a better in-sample fit, and this is especially true for housing starts for the first half of the sample. The difference in forecast errors is also small, but using revised lags leads to at least as accurate forecasts as the first release data (Table 2.10).

Table 2.9: Estimation results of the GARCH-MIDAS-X model: revised vs. first release data

Variable (X)	μ	α	β	γ	θ	ω_1	ω_2	m	LLF	VR
Industrial production										
Revised	0.0463***	0.0189***	0.8987***	0.1127***	-0.0452***	1	4.5673***	-0.0878	-14317.10	6.01
	(0.0073)	(0.0052)	(0.0141)	(0.0184)	(0.0123)		(1.2065)	(0.0949)		
First release	0.0463***	0.0185***	0.8991***	0.1129***	-0.0547***	1	4.7147***	-0.0943	-14315.58	6.56
	(0.0073)	(0.0052)	(0.0140)	(0.0184)	(0.0146)		(1.1927)	(0.0936)		
Housing starts										
Revised	0.0470***	0.0189***	0.8948***	0.1159***	-0.0173***	3.0722	4.8291*	-0.0954	-14299.24	16.59
	(0.0073)	(0.0053)	(0.0144)	(0.0185)	(0.0037)	(1.2979)	(2.1952)	(0.0899)		
First release	0.0475***	0.0210***	0.8959***	0.1128***	-0.0166***	3.0587	5.2011	-0.1082	-14307.50	14.57
	(0.0073)	(0.0053)	(0.0141)	(0.0181)	(0.0050)	(1.6837)	(2.9369)	(0.0952)		

*Bollerslev-Wooldridge QMLE robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. For the weight parameters I test $\omega_i = 1$. A "1" in the table indicates ω_i is fixed to 1 before estimation. LLF is the value of the log-likelihood function and VR is the variance ratio from Section 2.3 multiplied by 100. The MIDAS polynomial: $\log \tau_t = m + \theta \sum_{k=1}^{K} \varphi_k(\omega_1, \omega_2) X_{t-k}$. Whether a restricted ($\omega_1 = 1$) or an unrestricted weighting scheme is used is decided based on the revised data.*

Table 2.10: Out-of-sample results: revised vs. first release data

	1 quarter ahead		2 quarters ahead		3 quarters ahead		4 quarters ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Industrial production	1.00	0.99	1.00	1.00	1.00	1.00	0.99	1.00
Housing starts	1.00	0.99	0.97	0.96	0.98	0.96	0.99	0.96

MAFE ratio: $\frac{MAFE_{REV}}{MAFE_{FR}}$, where REV stands for the GARCH-MIDAS model driven by the data set where lags are revised and FR to the one where first release data is used. MSFE ratio calculated equivalently. A value below one means that the model using revised lags outperforms the model using first release data.

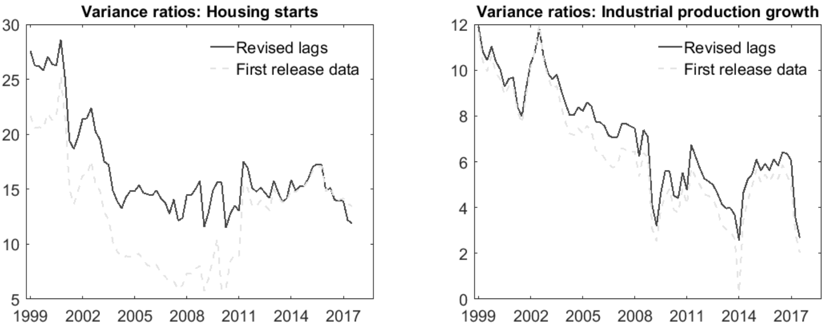


Figure 2.4: Variance ratios for models with revised lags and first release data.

2.C Controlling for realised volatility

For macroeconomic data to be a useful predictor of volatility it needs to contain information in addition to that in past realised volatility (RV). In this appendix I check the robustness of my in-sample results (Section 2.5) to including realised volatility (measured as the sum of the absolute value of returns) in the MIDAS polynomial. For each variable I keep the earlier choice of a restricted or unrestricted weighting scheme but re-estimate the weight parameter(s).

As expected, the variance ratios in Table 2.11 are higher than those in Table 2.3. The highest variance ratio (40.15) is achieved for the model augmented by the third principal component, highlighting the benefits of combining information from many variables. Overall, the models now explain roughly 35% of the total variance of stock returns.[37]

Although the in-sample results indicate that combining realised volatility with macroeconomic variables is useful for modelling long-term volatility, Conrad and Loch (2014) concluded that forecasting performance of these models resemble that of the model only driven by realised volatility. Hence I do not include models combining realised volatility and macroeconomic variables in the out-of-sample analysis.

When realised volatility is included as the third variable in a GARCH-MIDAS model (Table 2.12) some of the weighting schemes change clearly from earlier. As a robustness check I have reestimated all models in Table 2.12, first fixing the weighting scheme for each variable to that in Table 2.3. There are only very slight differences in the significance of the estimates for θ, and the main conclusions remain robust. Results are available upon request.

[37]The results are robust to using the sum of squared returns as the measure of realised volatility. The results are available upon request.

Table 2.11: Estimation results of the GARCH-MIDAS-RV-X model

Variable (X)	θ_{RV}	θ	ω_{12}	ω_{21}	ω_{22}	m	LLF	VR
Cons. conf.	0.0176***	-0.0822***	11.6222***	1.6150	3.0362	-1.0136***	-14265.61	33.92
	(0.0023)	(0.0298)	(3.1991)	(0.7056)	(1.9247)	(0.1268)		
News Heard	0.0173***	-0.0429***	11.7876***	1.5464*	1.6770**	-1.0048***	-14260.58	35.29
	(0.0020)	(0.0100)	(3.1963)	(0.2998)	(0.3180)	(0.1151)		
Buying cond.	0.0163***	-0.0668***	11.1347***	1	2.5344*	-0.9546***	-14257.88	36.14
	(0.0026)	(0.0207)	(3.5588)		(0.8683)	(0.1397)		
ISM NO	0.0173***	-0.0248***	12.3366***	1	5.6654	0.3466	-14263.97	33.52
	(0.0022)	(0.0082)	(3.1990)		(2.9875)	(0.5267)		
ISM RI	0.0182***	-0.0335***	11.4937***	1	2.4262**	-0.7673***	-14266.47	32.53
	(0.0020)	(0.0103)	(2.6927)		(0.5738)	(0.1731)		
SPF 1Q ahead	0.0172***	0.0084***	10.5520***	1	130.2367***	-1.1620***	-14263.82	32.03
	(0.0020)	(0.0019)	(2.9307)		(21.2599)	(0.1042)		
SPF 4Q ahead	0.0196***	0.0417***	11.0552***	2.2655	1.9912	-1.8449***	-14251.07	35.69
	(0.0018)	(0.0114)	(2.1781)	(2.2730)	(1.0515)	(0.1996)		
Ind. prod.	0.0187***	-0.0176*	12.5420***	1	5.9875**	-1.0319***	-14272.53	31.52
	(0.0021)	(0.0095)	(2.9226)		(2.0741)	(0.1284)		
ADS index	0.0178***	-0.1552**	13.2753***	1	8.2347*	-1.0494***	-14270.44	32.03
	(0.0024)	(0.0783)	(3.5753)		(4.2623)	(0.1232)		
Housing starts	0.0172***	-0.0082***	13.4911***	4.4822	8.8412	-0.9600***	-14261.35	35.65
	(0.0022)	(0.0028)	(2.9342)	(2.2209)	(5.2577)	(0.1331)		
Term spread	0.0194***	-0.1638***	10.3631***	4.9026	9.2134	-0.8243***	-14253.33	37.83
	(0.0018)	(0.0331)	(2.3620)	(3.7301)	(5.9166)	(0.1261)		
PC 1	0.0158***	0.1233***	13.3388***	1	3.7360***	-0.9432***	-14259.58	34.57
	(0.0024)	(0.0341)	(3.7809)		(1.1305)	(0.1255)		
PC 2	0.0188***	0.2836**	13.7891***	2.6058	2.7994	-1.0809***	-14252.77	37.20
	(0.0023)	(0.0799)	(4.0789)	(1.4803)	(1.3587)	(0.1237)		
PC 3	0.0197***	0.2978***	9.7658***	6.0263**	8.5303**	-1.1306***	-14239.38	40.15
	(0.0019)	(0.0511)	(2.1976)	(2.1682)	(3.4022)	(0.1053)		

*Bollerslev-Wooldridge QMLE robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. For the weight parameters I test $\omega_i = 1$. A "1" in the table indicates ω_i is fixed to 1 before estimation. LLF is the value of the log-likelihood function and VR is the variance ratio from Section 2.3 multiplied by 100. The MIDAS polynomial: $\log \tau_t = m + \theta_{RV} \sum_{k=1}^{K} \varphi_k(1, \omega_{12}) RV_{t-k} + \theta \sum_{k=1}^{K} \varphi_k(\omega_{21}, \omega_{22}) X_{t-k}$. $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$. See notes on Table 2.1 for full names of explanatory variables.*

Table 2.12: Estimation results of the GARCH-MIDAS-RV-X_1-X_2 model

Variables ($X_1 + X_2$)	θ_{RV}	θ_1	θ_2	ω_{12}	ω_{21}	ω_{22}	ω_{31}	ω_{32}	m	LLF	VR
IP + ISM NO	0.0204***	0.0765***	-0.0500***	9.0755***	1	1.0000	1	2.1580	1.4006*	-14253.53	36.28
	(0.0027)	(0.0253)	(0.0153)	(2.9311)		(0.4269)		(1.0419)	(0.8198)		
IP + ISM RI	0.0182***	-0.0010	-0.0329***	11.5865***	1	4.4890**	1	2.4337**	-0.7670***	-14266.46	32.54
	(0.0023)	(0.0119)	(0.0119)	(3.0626)		(1.5836)		(0.5648)	(0.1733)		
IP + Cons. conf.	0.0173***	-0.0046	-0.0770**	12.1434***	1	5.3251	1.5623	3.0084	-0.9892***	-14265.88	33.96
	(0.0024)	(0.0116)	(0.0335)	(3.6318)		(2.7796)	(0.7495)	(2.0866)	(0.1390)		
IP + News Heard	0.0192***	0.0288	-0.0400*	10.0913**	1	1.0000	1.5886	1.4679	-1.1588***	-14259.13	35.50
	(0.0036)	(0.0398)	(0.0234)	(3.5593)		(1.6035)	(0.4156)	(0.6346)	(0.2569)		
IP + Buying cond.	0.0180***	0.0493*	-0.0870***	9.2418***	1	1.0000	1	1.5927	-1.1453***	-14252.75	37.31
	(0.0027)	(0.0269)	(0.0253)	(3.1643)		(0.5923)		(0.4574)	(0.1662)		
IP + SPF 1Q ahead	0.0170***	-0.0027	0.0081***	10.8548***	1	4.2412**	1	130.1162***	-1.1424***	-14263.78	32.11
	(0.0023)	(0.0127)	(0.0024)	(3.5935)		(1.3427)		(15.7655)	(0.1504)		
IP + SPF 4Q ahead	0.0181***	-0.0166	0.0421***	12.4859***	1	4.5235	1.3175	1.4658	-1.7473***	-14249.74	36.15
	(0.0022)	(0.0143)	(0.0118)	(3.0050)		(2.2235)	(1.6521)	(0.9141)	(0.2097)		
IP + Term spread	0.0211***	0.0358	-0.1656***	8.3364***	1	1.0000	7.0881	10.3478	-0.9830***	-14250.06	38.81
	(0.0024)	(0.0218)	(0.0450)	(2.7990)		(0.8057)	(8.9913)	(6.6892)	(0.1795)		
ADS + ISM NO	0.0172***	-0.0125	-0.0241***	12.5629***	1	50.3635	1	5.3465	0.3100	-14263.93	33.53
	(0.0022)	(0.0706)	(0.0091)	(3.4956)		(66.1549)		(3.5904)	(0.5654)		
ADS + ISM RI	0.0180***	-0.0258	-0.0302**	11.8089***	1	96.5555***	1	2.3495**	-0.7876***	-14266.28	32.58
	(0.0021)	(0.0515)	(0.0119)	(2.9473)		(25.2306)		(0.5829)	(0.1745)		
ADS + Cons. conf.	0.0172***	-0.0395	-0.0714**	12.3561***	1	103.4087***	1.7330	3.2365	-1.0001***	-14265.54	33.94
	(0.0025)	(0.0552)	(0.0349)	(3.8112)		(21.1130)	(1.0057)	(2.9075)	(0.1299)		
ADS + News Heard	0.0168***	-0.0347	-0.0397***	12.6078***	1	94.9688***	1.5902*	1.6767**	-0.9872***	-14260.60	35.27
	(0.0022)	(0.0567)	(0.0116)	(3.9045)		(21.0470)	(0.3440)	(0.3341)	(0.1203)		
ADS + Buying cond.	0.0165***	0.0254	-0.0710***	10.7730***	1	14.8346	1	2.5418*	-0.9607***	-14257.75	36.26
	(0.0028)	(0.0714)	(0.0228)	(3.7483)		(14.7636)		(0.8405)	(0.1463)		
ADS + SPF 1Q ahead	0.0172***	0.0000	0.0084***	10.5522***	1	5.8810	1	215.7699***	-1.1621***	-14263.82	32.04
	(0.0020)	(0.142)	(0.0020)	(3.0005)		(107.1975)		(40.6125)	(0.1080)		
ADS + SPF 4Q ahead	0.0173***	-0.1482	0.0419***	13.2065***	1	6.0912	1.1523	1.3326	-1.7629***	-14248.27	36.48
	(0.0026)	(0.1176)	(0.0119)	(3.6218)		(4.4823)	(1.5551)	(0.8500)	(0.1998)		
ADS + Term spread	0.0188***	-0.0269	-0.1570***	10.8196***	1	135.5776***	5.1801	9.5113	-0.8134***	-14253.39	37.79
	(0.0021)	(0.0450)	(0.0360)	(2.7323)		(24.1223)	(5.1998)	(7.9265)	(0.1280)		

Table 2.12 *continued*

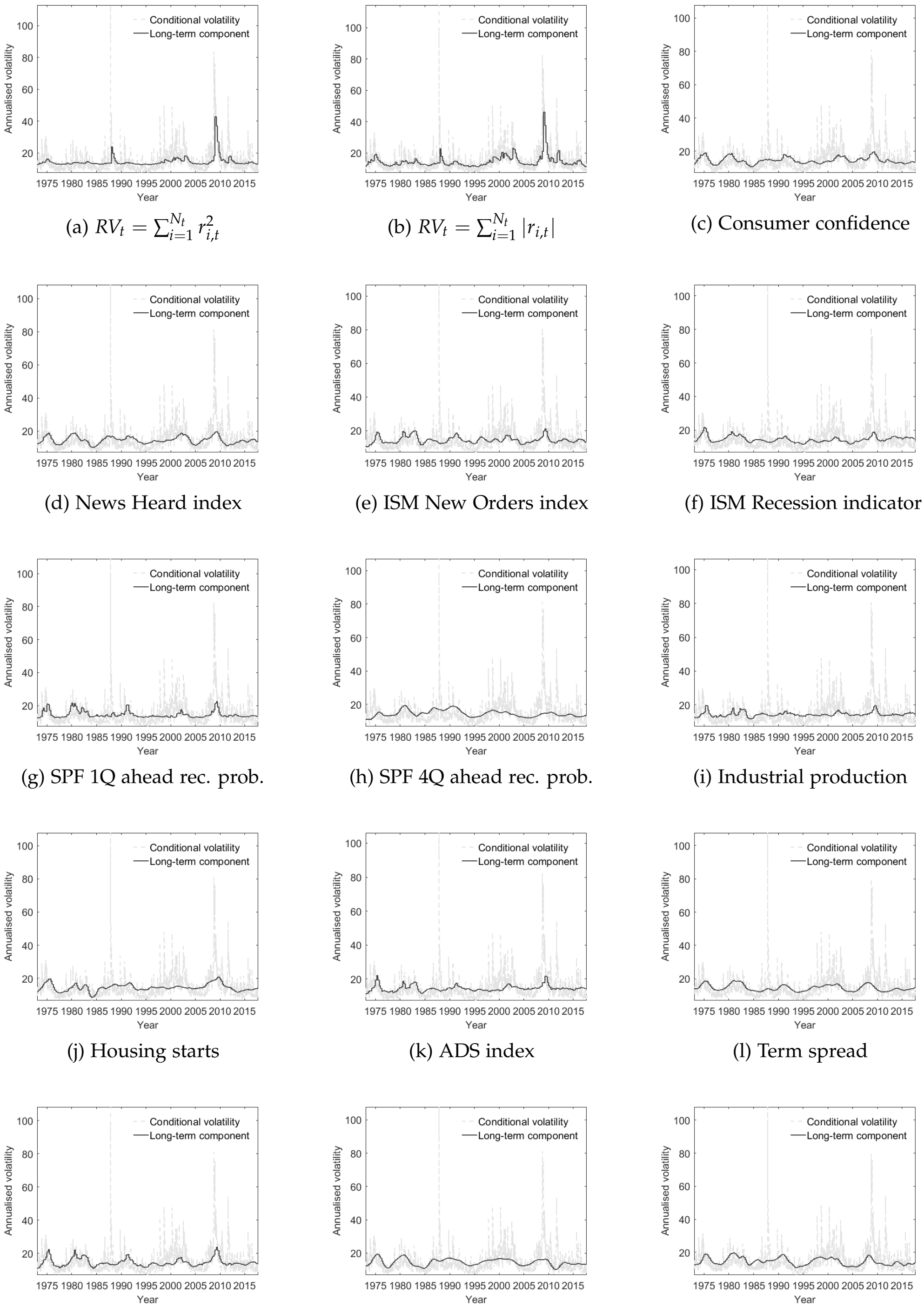
Variables ($X_1 + X_2$)	θ_{RV}	θ_1	θ_2	ω_{12}	ω_{21}	ω_{22}	ω_{31}	ω_{32}	m	LLF	VR
HS + ISM NO	0.0160***	-0.0062*	-0.0205**	15.0784***	5.0932	11.8092	1	2.9621	0.2081	-14256.95	36.38
	(0.0024)	(0.0090)	(0.0034)	(3.6691)	(5.5068)	(15.7084)		(1.5624)	(0.5451)		
HS + ISM RI	0.0162***	-0.0066**	-0.0228**	13.8936***	4.8515	9.5505	1	2.1647	-0.7312***	-14258.07	35.86
	(0.0024)	(0.0028)	(0.0108)	(3.3155)	(2.6529)	(6.9089)		(0.7972)	(0.1782)		
HS + Cons. conf.	0.0169***	-0.0076**	-0.0182	12.9488***	5.0501*	9.1930	6.6663	35.9247	-0.9492***	-14259.27	36.54
	(0.0023)	(0.0031)	(0.0236)	(3.5927)	(2.3921)	(5.5497)	(13.3102)	(91.1851)	(0.1367)		
HS + News Heard	0.0167***	-0.0086***	-0.0057	13.3428***	4.6220	7.7150	13.4874	61.7484	-0.9336***	-14257.69	37.02
	(0.0024)	(0.0031)	(0.0064)	(3.9802)	(2.3671)	(5.0564)	(39.2856)	(186.2468)	(0.1429)		
HS + Buying cond.	0.0168***	-0.0053	-0.0311	11.0114***	6.2412	12.0504	1	5.3371	-0.9555***	-14254.56	37.75
	(0.0027)	(0.0039)	(0.0272)	(3.1731)	(3.8425)	(11.4658)		(6.2952)	(0.1472)		
HS + SPF 1Q ahead	0.0154***	-0.0061**	0.0060***	13.1431***	5.6021**	10.4137*	1	120.6965***	-1.0043***	-14256.50	35.31
	(0.0023)	(0.0026)	(0.0022)	(3.5610)	(2.3417)	(5.4420)		(8.5692)	(0.1321)		
HS + SPF 4Q ahead	0.0186***	-0.0027**	0.0375***	12.5380***	80.0777***	237.5071***	2.5138	2.0776	-1.7119***	-14242.03	37.31
	(0.0018)	(0.0011)	(0.0113)	(2.6024)	(23.7847)	(89.2930)	(2.7471)	(1.2024)	(0.2056)		
HS + Term spread	0.0181***	-0.0025**	-0.1532***	12.0082***	99.3805***	293.9653**	3.5422	5.7496	-0.7739***	-14245.25	39.33
	(0.0018)	(0.0011)	(0.0432)	(2.8959)	(32.4306)	(129.9762)	(8.6210)	(11.8705)	(0.1364)		
TS + ISM NO	0.0185***	-0.1447***	-0.0068	10.6966***	4.6840	8.7854	1	8.4334	-0.4485	-14252.88	37.68
	(0.0022)	(0.0424)	(0.0102)	(2.9709)	(5.5924)	(7.8974)		(15.7129)	(0.5929)		
TS + ISM RI	0.0192***	-0.1385***	-0.0051	9.7491***	133.9744	235.4417*	1	71.7352	-0.8157***	-14253.14	37.91
	(0.0017)	(0.0282)	(0.0057)	(2.6071)	(81.2545)	(134.5007)		(52.6096)	(0.1308)		
TS + Cons. conf.	0.0184***	-0.1564***	-0.0258	9.7518**	6.3073	9.6275	4.5944	21.8797	-0.7928***	-14250.31	39.10
	(0.0027)	(0.0418)	(0.0395)	(3.0940)	(9.5595)	(11.7942)	(12.5499)	(82.0642)	(0.1371)		
TS + News Heard	0.0178***	-0.1319***	-0.0210	10.6887***	5.2582	8.6007	1.3828	1.7296	-0.8051***	-14250.96	38.74
	(0.0021)	(0.0411)	(0.0128)	(3.2264)	(6.5670)	(8.4118)	(0.4753)	(0.6726)	(0.1287)		
TS + Buying cond.	0.0187***	-0.1171***	-0.0227**	8.8588***	166.4777	291.7320	1	6.8949*	-0.8730***	-14248.12	39.15
	(0.0020)	(0.0302)	(0.0101)	(2.5852)	(135.6494)	(200.0857)		(3.1793)	(0.1221)		
TS + SPF 1Q ahead	0.0184***	-0.1436***	0.0026	10.4301***	4.9557	9.1253	1	156.1596***	-0.8626***	-14252.86	37.27
	(0.0020)	(0.0423)	(0.0025)	(2.6560)	(5.2759)	(7.5944)		(14.4404)	(0.1368)		
TS + SPF 4Q ahead	0.0188***	-0.1276***	0.0333***	10.4753***	8.1591	12.9019	1.0000	1.3741	-1.4472***	-14239.95	39.94
	(0.0018)	(0.0404)	(0.0116)	(2.6456)	(14.1907)	(19.2513)	(1.3330)	(0.8008)	(0.2410)		
PC 1 + PC 2	0.0161***	0.0815***	0.1077***	17.0762***	1	8.8813**	300	183.9172	-0.9589***	-14247.82	37.50
	(0.0022)	(0.0275)	(0.0278)	(4.3941)		(3.4994)	(68.6996)	(39.6739)	(0.1209)		
PC 1 + PC 3	0.0192***	0.0135	0.2828***	9.8368***	1	7.9971	6.1073**	8.5988*	-1.1058***	-14239.60	40.11
	(0.0024)	(0.0275)	(0.0592)	(2.5311)		(16.3626)	(2.3641)	(3.8182)	(0.1264)		
PC 2 + PC 3	0.0192***	0.0576**	0.2537***	14.2060***	300***	181.8189***	7.5564***	16.7958***	-1.1062***	-14236.19	40.91
	(0.0019)	(0.0273)	(0.0573)	(3.0446)	(61.5563)	(30.9520)	(2.6088)	(6.1767)	(0.1061)		

*Bollerslev-Wooldridge QMLE robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. For the weight parameters I test $\omega_i = 1$. A "1" in the table indicates ω_i is fixed to 1 before estimation. LLF is the value of the log-likelihood function and VR is the variance ratio from Section 2.3 multiplied by 100. $\omega_{11} = 1$ in all specifications. The MIDAS polynomial: $\log \tau_t = m + \theta_{RV} \sum_{k=1}^{K} \varphi_k(1, \omega_{12}) RV_{t-k} + \theta_1 \sum_{k=1}^{K} \varphi_k(\omega_{21}, \omega_{22}) X_{1,t-k} + \theta_2 \sum_{k=1}^{K} \varphi_k(\omega_{31}, \omega_{32}) X_{2,t-k}$, where X_1 denotes the macroeconomic data and X_2 the sentiment data, as stated in the first column. See notes on Table 2.1 for full names of explanatory variables.*

2.D Figures of baseline results

Figure 2.5 shows how the GARCH-MIDAS-X model decomposes volatility into two components by drawing total volatility and the extracted long-term component separately based on the models identified in Table 2.3 in Section 2.5.1. It is clear that the long-term components based on different variables capture long-term volatility in very different ways. For example, the realised volatility measures capture the peaks while the macroeconomic and sentiment data capture the low-frequency movements in volatility well.

The weights of the lags of the macroeconomic and confidence variables are shown in Figure 2.6. As expected, realised volatility, industrial production and the ADS index have decaying weighting schemes. This is intuitive: recent information in realised volatility or variables describing the current economic situation is more important than older information. For the consumer confidence indicators the optimal weighting schemes are often hump-shaped, indicating that older information is more important for volatility than very recent information. Thus, business confidence, which gets decaying weights, seems to anticipate stock market volatility less than consumer confidence. The largest weight for housing starts is on the fourth and fifth lags, which supports the perception that housing starts is a forward-looking indicator.



(a) $RV_t = \sum_{i=1}^{N_t} r_{i,t}^2$

(b) $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$

(c) Consumer confidence

(d) News Heard index

(e) ISM New Orders index

(f) ISM Recession indicator

(g) SPF 1Q ahead rec. prob.

(h) SPF 4Q ahead rec. prob.

(i) Industrial production

(j) Housing starts

(k) ADS index

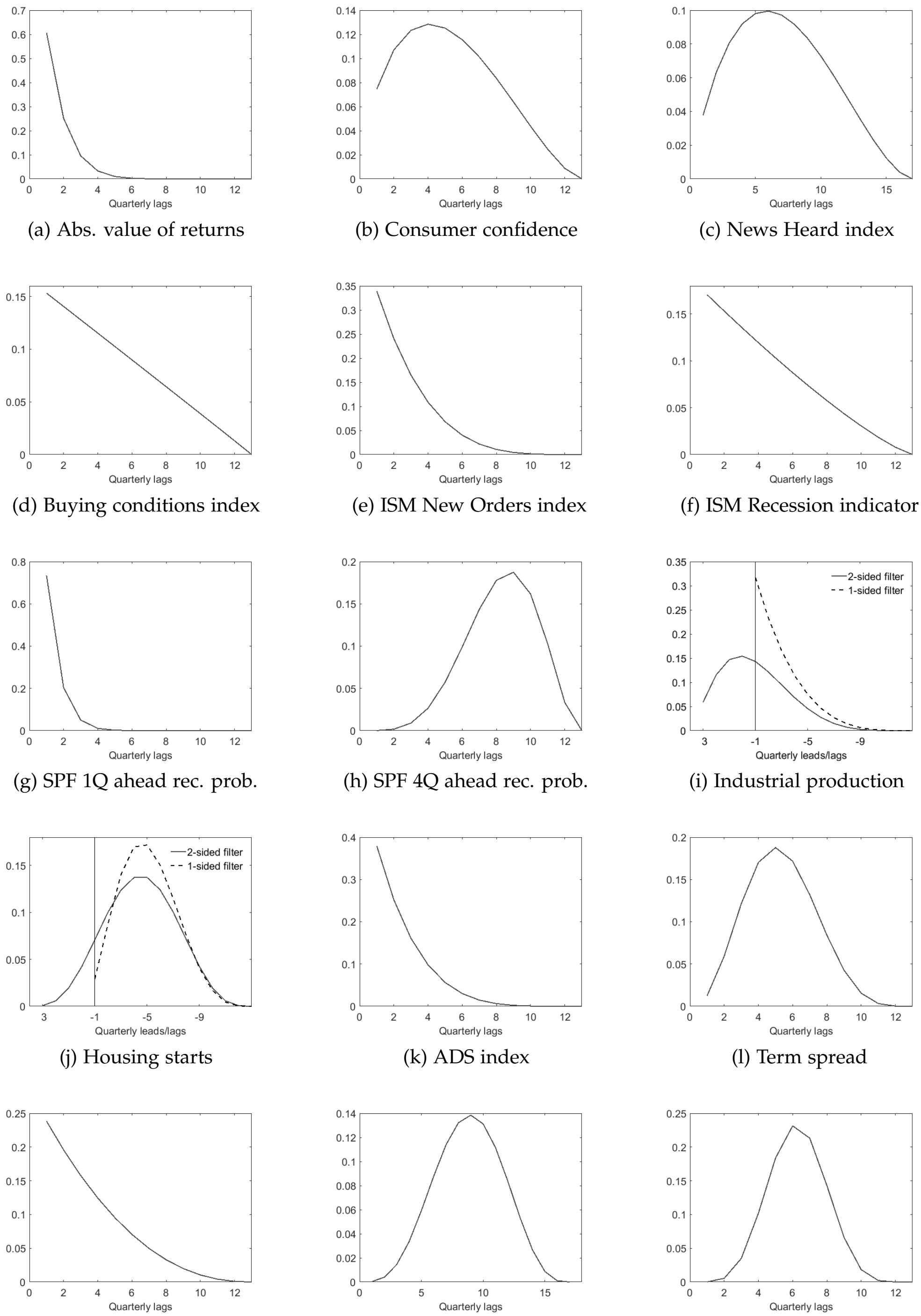
(l) Term spread

(m) Principal component 1

(n) Principal component 2

(o) Principal component 3

Figure 2.5: Total ($\tau_t g_{i,t}$) and long-term (τ_t) volatility (annualised) of selected GARCH-MIDAS-X models from Table 2.3



(a) Abs. value of returns
(b) Consumer confidence
(c) News Heard index
(d) Buying conditions index
(e) ISM New Orders index
(f) ISM Recession indicator
(g) SPF 1Q ahead rec. prob.
(h) SPF 4Q ahead rec. prob.
(i) Industrial production
(j) Housing starts
(k) ADS index
(l) Term spread
(m) Principal component 1
(n) Principal component 2
(o) Principal component 3

Figure 2.6: MIDAS weight profiles of selected GARCH-MIDAS-X models from Table 2.3

2.E One-sided and two-sided MIDAS filters

This appendix compares the results for using a one-sided and a feasible two-sided MIDAS filter for industrial production growth and housing starts. The two-sided filters are based on the Survey of Professional Forecasters' (SPF) forecasts. The SPF forecasts range from nowcasts to four quarter ahead forecasts, and I augment my dataset using all these forecast horizons together with twelve lags as before.

Table 2.3 presented the in-sample results using both one-sided and two-sided filters, showing that the in-sample fit, as measured by the variance ratio, improves for the two-sided filters, in line with the results in Conrad and Loch (2014). Looking at the variance ratios calculated over rolling windows in Figure 2.7, the improvement in in-sample fit is especially large for industrial production. Table 2.13 compares the out-of-sample forecast errors for the one-sided and two-sided filters. The difference in forecasting performance is, however, small for both variables, which is in line with the MSE ratios reported in Conrad and Loch (2014). The largest difference is for the one quarter ahead forecast when using industrial production growth, for which the two-sided filter gives a clearly smaller forecast error.

When combining macroeconomic and sentiment data we can see that the in-sample fit mostly improves in particular when using the two-sided filter for industrial production growth (Table 2.15). The significance of the industrial production index strengthens somewhat, but in-sample fit improves only marginally compared to models where industrial production is allowed to have an unrestricted weighting scheme (but no leads).[38] For housing starts the gain in in-sample fit is relatively small and the model parameters are very similarly estimated regardless of whether SPF data is used or not.

Table 2.13: Out-of-sample results: one-sided vs. two-sided MIDAS filters

	1 quarter ahead		2 quarters ahead		3 quarters ahead		4 quarters ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Industrial production	0.98	0.88	1.01	0.99	1.02	0.99	0.99	0.99
Housing starts	0.99	1.03	0.99	1.00	0.99	0.99	0.98	0.99

MAFE ratio: $\frac{MAFE_{2s}}{MAFE_{1s}}$, where 2s stands for the GARCH-MIDAS model with a 2-sided MIDAS filter and 1s to the one where a 1-sided filter is used. MSFE ratio calculated equivalently. A value below one means that the model using the 2-sided filter outperforms the model using the 1-sided filter.

[38]Results are available upon request.

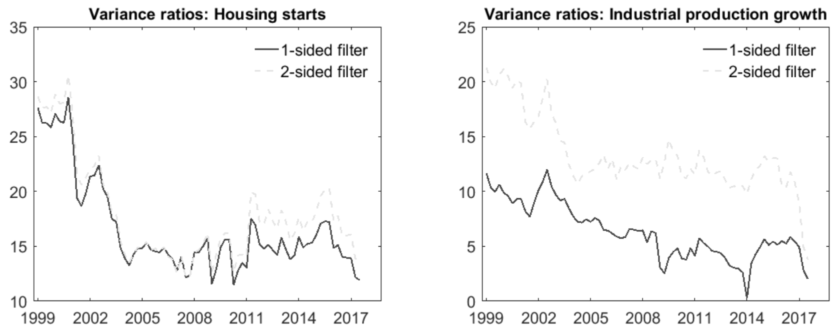


Figure 2.7: Variance ratios for models with 2-sided and 1-sided MIDAS filters.

Table 2.14: Out-of-sample results: GARCH-MIDAS-IP-X_2 model vs. GARCH-MIDAS-X model

Benchmark: GARCH-MIDAS-X model, where X is industrial production

	MAFE ratio				MSFE ratio			
Variable	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
Consumer confidence	0.99	0.99	0.99	1.00	0.92	0.97	1.00	1.00
News Heard index	0.97	0.93***	0.90***	0.92***	1.04	0.96	0.95	0.96
Buying conditions	0.97	0.93**	0.91***	0.95	1.07	0.88	0.91	0.94
ISM New Orders index	0.97*	0.97	0.97	0.96	1.01	1.00	1.00	1.00
ISM Recession indicator	1.05	1.08	1.07	1.07	0.92	1.00	1.00	1.01
SPF 1Q ahead recession probability	0.94*	0.98	1.00	1.00	0.83	0.99	0.98	1.00
SPF 4Q ahead recession probability	1.01	0.99	0.95	0.95	1.12	1.03**	1.01	1.02
Term spread	1.02	0.93**	0.89***	0.86***	1.34	0.96	0.92	0.92

Benchmark: GARCH-MIDAS-X model, where X is sentiment data (as indicated in the first column)

	MAFE ratio				MSFE ratio			
Variable	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
Consumer confidence	1.05**	1.03*	1.04***	1.03***	1.03	1.04	1.04	1.02
News Heard index	1.02	0.99	0.98	1.00	1.04	1.00	1.00	1.00
Buying conditions	1.03*	1.01	0.99	1.01	0.98	1.00	1.00	1.01
ISM New Orders index	1.01	1.00	1.00	0.99	1.06	1.00	1.00	1.00
ISM Recession indicator	1.02	1.03**	1.02	1.01	1.00	1.02*	1.01	1.00
SPF 1Q ahead recession probability	1.00	1.00	1.00	1.00	1.02	1.00	1.00	1.00
SPF 4Q ahead recession probability	1.02	1.01	1.00	1.00	1.05	0.98	0.97	0.98
Term spread	1.00	1.03	1.03	1.02	0.99	0.99	1.00	1.00

One-sided filter used for industrial production growth (IP). Top panel: $\frac{MAFE_{IP+sentiment}}{MAFE_{IP}}$. *Bottom panel:* $\frac{MAFE_{IP+sentiment}}{MAFE_{sentiment}}$. *MSFE ratios calculated equivalently. A value below 1 means the model combining IP and sentiment data in the long-term component outperforms the model driven by only IP or sentiment data. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5% and 1% level, respectively, according to the Giacomini and White (2006) test.*

Table 2.15: Estimation results of the GARCH-MIDAS-X_1-X_2 model using one or two-sided filters

Variables ($X_1 + X_2$)	θ_1	θ_2	ω_{11}	ω_{12}	ω_{21}	ω_{22}	m	LLF	VR
IP + ISM New Orders index	0.0141	-0.0561***	1	9.9739	1	5.4146**	2.8456***	-14302.01	13.40
	(0.0155)	(0.0132)		(6.3460)		(2.0335)	(0.7048)		
IP + ISM Recession indicator	-0.0199	-0.0563***	1	2.8490***	1	2.4631***	0.3257**	-14305.48	10.95
	(0.0160)	(0.0143)		(0.7133)		(0.5333)	(0.1342)		
IP + Consumer confidence	-0.0183	-0.1440***	1	4.2088	1.6103	2.5086**	-0.1429	-14303.26	13.23
	(0.0165)	(0.0380)		(2.1567)	(0.4179)	(0.6487)	(0.0907)		
IP + News Heard index	-0.0083	-0.0821***	1	5.8774	1.8880***	2.6936***	-0.1715**	-14298.85	16.38
	(0.0161)	(0.0179)		(7.3722)	(0.3372)	(0.5973)	(0.0868)		
IP + Buying conditions index	-0.0074	-0.1176***	1	4.4513*	1	2.0122***	-0.1644**	-14292.88	17.81
	(0.0143)	(0.0249)		(1.8199)		(0.2856)	(0.0802)		
IP + SPF 1Q ahead	-0.0233	0.0133***	1	3.0314***	1	136.5962***	-0.4006***	-14298.39	13.34
	(0.0142)	(0.0024)		(0.7532)		(34.6525)	(0.1022)		
IP + SPF 4Q ahead	-0.0499**	0.0591***	1	2.9196**	1.0000	1.2036	-1.1230***	-14289.85	16.40
	(0.0241)	(0.0159)		(0.9083)	(1.5700)	(0.9225)	(0.2709)		
IP + Term spread	-0.0239	-0.2240***	1	5.2417**	1.5648	2.4056	0.2426	-14301.42	13.49
	(0.0198)	(0.0636)		(1.7899)	(4.2703)	(5.6387)	(0.1540)		
HS + ISM New Orders index	-0.0231***	-0.0547***	2.2646	3.5372*	1	2.3130	2.9187***	-14285.19	22.53
	(0.0059)	(0.0165)	(0.7678)	(1.3763)		(0.9219)	(0.9125)		
HS + ISM Recession indicator	-0.0296***	-0.0788***	1.0000	1.4906	1	1.7707*	0.6385**	-14283.94	22.11
	(0.0082)	(0.0243)	(0.6514)	(0.7219)		(0.4502)	(0.2599)		
HS + Consumer confidence	-0.0181**	-0.0419	5.3350	4.5580	4.0484	16.5744	-0.0949	-14293.78	20.25
	(0.0087)	(0.0825)	(7.1738)	(6.2043)	(10.2182)	(64.2043)	(0.0996)		
HS + News Heard index	-0.0178***	-0.0102	6.2949	6.6214	30.5774	10.1873	-0.1000	-14292.88	17.98
	(0.0062)	(0.0065)	(4.9008)	(4.2059)	(33.9743)	(9.2695)	(0.0923)		
HS + Buying conditions index	-0.0085*	-0.0848***	9.7124	9.0372	1	2.1381**	-0.1428*	-14287.81	21.44
	(0.0045)	(0.0245)	(8.7730)	(7.6191)		(0.5023)	(0.0822)		
HS + SPF 1Q ahead	-0.0110***	0.0113***	11.4105	9.4720	1	138.0426***	-0.3603***	-14284.99	20.13
	(0.0040)	(0.0027)	(8.5380)	(6.7198)		(12.9747)	(0.1079)		
HS + SPF 4Q ahead	-0.0184***	0.0339***	4.1701*	4.9430**	80.5261	20.5261	-0.6944***	-14280.41	21.13
	(0.0050)	(0.0096)	(1.4264)	(1.4466)	(131.6139)	(30.1964)	(0.1841)		
HS + Term spread	-0.0207***	-0.2207***	3.0712	4.0861**	1.6227	1.1337	0.2871**	-14285.73	23.88
	(0.0057)	(0.0632)	(1.3555)	(1.3295)	(3.5621)	(2.9447)	(0.1410)		

*Bollerslev-Wooldridge QMLE robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. For the weight parameters I test $\omega_i = 1$. A "1" in the table indicates ω_i is fixed to 1 before estimation. LLF is the value of the log-likelihood function and VR is the variance ratio from Section 2.3 multiplied by 100. The MIDAS polynomial: $\log \tau_t = m + \theta_1 \sum_{k=1}^{K} \varphi_k(\omega_{11}, \omega_{12}) X_{1,t-k} + \theta_2 \sum_{k=1}^{K} \varphi_k(\omega_{21}, \omega_{22}) X_{2,t-k}$, where X_1 denotes the macroeconomic data and X_2 the sentiment data as stated in the first column. For industrial production (IP) the one-sided filter is used, while the two-sided filter is used for housing starts (HS).*

2.F Excluding the financial crisis

The differences in the out-of-sample loss functions are very large in connection to the financial crisis, especially when squared forecast errors are considered. This appendix presents a robustness check to the MSFE ratios by excluding the periods during the financial crisis in which the forecast errors are huge. This of course changes the interpretation of the MSFE ratios in the sense that it gives no weight to the models' forecasting performance during these periods. Table 2.16 shows that the exclusion of the financial crisis clearly improves the performance of the realised volatility driven models over the shorter horizons. Almost all the other GARCH-MIDAS-X models are now worse than the GJR-GARCH(1,1) model for one quarter ahead forecasts. On the other hand, although the GARCH-MIDAS-X models seem to forecast slightly worse than before, over longer horizons the ranking of the models does not seem significantly altered. Comparing the performance of GARCH-MIDAS-X_1-X_2 models and GARCH-MIDAS-X models excluding the financial crisis we can see that the results in Table 2.17 are similar to those in Table 2.6, but statistical significance is stronger when the financial crisis is excluded.

Table 2.16: MSFE ratios excluding the financial crisis

Variable (X)	1Q ahead	2Q ahead	3Q ahead	4Q ahead
Squared returns	0.95	0.93	1.08	1.84
Absolute value of returns	0.93	0.92	1.03	1.63
Consumer confidence	1.04	0.99	0.95	0.94
News Heard index	1.03	0.98	0.96	0.95
Buying Conditions	1.03	0.99	0.96	0.93
ISM New Orders index	1.05	1.03	1.00	0.97
ISM Recession indicator	1.09*	1.11**	1.08	1.07
SPF 1Q ahead recession probability	1.07	1.05	1.02	1.03
SPF 4Q ahead recession probability	1.10*	1.07	1.03	1.00
Industrial production	1.05*	1.02	1.00	0.98
ADS index	1.01	0.97	0.97	0.95
Housing starts	0.98	0.93**	0.94*	0.93*
Term spread	1.10	1.00	0.94	0.90
Principal component 1	1.02	1.00	0.99	0.98
Principal component 2	1.08	1.07	1.03	1.01
Principal component 3	1.16*	1.06	1.01	0.96

Benchmark model: GJR-GARCH(1,1). MSFE ratio: $\frac{MSFE_{GMX}}{MSFE_{GARCH}}$*, where GMX stands for the GARCH-MIDAS-X model. A value below 1 means the GARCH-MIDAS-X model outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5% and 1% level, respectively, according to the Giacomini and White (2006) test. The large peak in the forecast errors during the financial crisis has been removed (Q3 2008 - Q4 2009).*

Table 2.17: MSFE ratios excluding the financial crisis: GARCH-MIDAS-X_1-X_2 model vs. GARCH-MIDAS-X model

Benchmark: GARCH-MIDAS-X model, where X is macroeconomic data (as indicated in the first row)

	Industrial production				ADS index				Housing starts				Term spread			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
CC	1.02	0.97	0.97	0.97	1.06	1.08	0.99	0.99	0.99	0.99	0.99	1.00	1.01	0.99	0.98	0.99
NH	1.00	0.96	0.95	0.95	1.03	1.01	0.97	0.98	1.02	0.99	0.95	0.97	0.99	0.98	0.97	0.97
BC	1.03	1.03	1.01	1.00	1.03	1.07	1.04	1.05	1.04	1.03	0.99	1.00	0.99	1.00	1.00	1.00
ISM NO	1.00	1.02	1.01	0.99	1.05	1.09	1.09	1.07	1.04	1.04	1.00	0.99	1.00	1.02	0.99	0.96
ISM RI	1.07*	1.13*	1.11	1.12	1.14***	1.17**	1.15*	1.18**	1.04	1.07***	1.04	1.05	1.01	0.99	0.95	0.99
SPF 1Q	1.03	1.03*	1.01	1.03	1.05	1.08**	1.03	1.06*	1.07*	1.08*	1.04	1.03	1.01	1.02*	1.03	1.03
SPF 4Q	1.10**	1.08*	1.02	1.02	1.10**	1.11**	1.07	1.06	1.13**	1.13**	1.05	1.04	1.07***	1.07**	1.08**	1.08*
TS	1.06	0.98	0.94	0.91	1.09	1.03	0.97	0.95	1.10	1.05	0.98	0.96				

Benchmark: GARCH-MIDAS-X model, where X is sentiment data (as indicated in the first column)

	Industrial production				ADS index				Housing starts				Term spread			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
CC	1.03	1.00	1.01	1.02	1.03	1.06	1.01	1.00	0.93*	0.93	0.98	1.00	1.07	0.99	0.97	0.95
NH	1.01	1.00	0.99	0.99	1.01	1.01	0.99	0.98	0.97	0.94	0.93	0.96	1.06	1.00	0.95*	0.93*
BC	1.05	1.05	1.04	1.05	1.01	1.05	1.05	1.06	0.99	0.96	0.97	1.00	1.06	1.00	0.98	0.96
ISM NO	1.00	1.00	1.01	1.01	1.02	1.03	1.06	1.05	0.97	0.94**	0.95	0.95	1.05	0.98	0.93**	0.89***
ISM RI	1.02	1.04*	1.02	1.04	1.05*	1.03	1.03	1.05	0.93***	0.90***	0.90*	0.92	1.02	0.89**	0.82***	0.84**
SPF 1Q	1.01	1.00	0.99	0.98	1.00	1.00	0.98	0.97	0.97	0.96**	0.96*	0.93***	1.04	0.97	0.94	0.90*
SPF 4Q	1.05**	1.03	0.99	1.01	1.02	1.01	1.01	1.00	1.00	0.98	0.96*	0.97	1.07	0.99	0.98	0.98
TS	1.01	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.98	0.98	0.99	0.99				

Benchmark: GARCH-MIDAS-X model, where X is a principal component (as indicated in the first column

	Principal component 1				Principal component 2				Principal component 3			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
PC 1					1.06	1.04	1.01	0.98	1.13	1.04	0.99	0.94
PC 2	1.00	0.97	0.97	0.96					1.11*	1.01	0.99	0.95
PC 3	0.99	0.98	0.97	0.97	1.03	1.02	1.00	1.00				

*Top panel: MSFE ratio of a GARCH-MIDAS-X_1-X_2 model including macroeconomic and sentiment data and a GARCH-MIDAS-X model including macroeconomic data: $\frac{MSFE_{macro+sentiment}}{MSFE_{macro}}$. Middle panel: MSFE ratio of a GARCH-MIDAS-X_1-X_2 model including macroeconomic and sentiment data and a GARCH-MIDAS-X model including sentiment data: $\frac{MSFE_{macro+sentiment}}{MSFE_{sentiment}}$. A value below 1 means the model combining macroeconomic and sentiment data outperforms the model driven by macroeconomic or sentiment data. Bottom panel: MSFE ratio of a GARCH-MIDAS-X_1-X_2 model including two principal components and a GARCH-MIDAS-X model, where X is the principal component corresponding to the first column: $\frac{MSFE_{PC_{row}+PC_{column}}}{MSFE_{PC_{row}}}$. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5% and 1% level, respectively, according to the Giacomini and White (2006) test. See notes on Table 2.1 for full names of explanatory variables. The large peak in the forecast errors during the financial crisis has been removed (Q3 2008 - Q4 2009).*

2.G Model Confidence Set procedure

This appendix presents a rough outline of the MCS procedure by Hansen et al. (2011) used in Section 2.6.1. The presentation follows closely Hansen et al. (2011), where details on the procedure can be found.

M^0 is the competing set of models or objects. M^* is the set of superior models, defined in Definition 1 in Hansen et al. (2011) as $M^* \equiv \{i \in M^0 : \mu_{ij} \leq 0, \forall j \in M^0\}$. The objective of the MCS procedure is to identify this superior set of models by eliminating the models that are significantly inferior to the other objects in M^0.

Relative performance is defined as $d_{ij,t} \equiv L_{i,t} - L_{j,t}$, for all $i, j \in M^0$, where L is a loss function and i and j index the models. $\mu_{ij} \equiv E(d_{ij,t})$ is assumed finite and independent of t. Models are ranked according to their expected loss, so that model i is preferred to model j if $\mu_{ij} < 0$. Hansen et al. (2011) define the t-statistics

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{var}(\bar{d}_{ij})}}, \quad for\ i, j \in M,$$

where $\bar{d}_{ij} \equiv \frac{1}{n}\Sigma_{t=1}^{n} d_{ij,t}$ and $\widehat{var}(\bar{d}_{ij})$ is the (bootstrapped) estimate of $var(\bar{d}_{ij})$. This t-statistic is associated with the null hypothesis $H_{ij} : \mu_{ij} = 0, \forall i, j \in M$ and the test statistic $T_{R,M} \equiv max|t_{i,j}|$. The corresponding coherent elimination rule is $e_{R,M} = arg\ max_{i \in M} sup_{j \in M} t_{ij}$. The asymptotic distribution of the test statistic is non-standard, and thus needs to be estimated using bootstrap.

The MCS algorithm (Definition 2 in Hansen et al. (2011)):
Step 0: Set $M = M^0$.
Step 1: Test $H_{0,M}$ using the equivalence test δ_M (i.e., test based on $T_{R,M}$) and significance level α. $H_{0,M} : \mu_{ij} = 0, \forall i, j \in M$.
Step 2: i) If δ_M is accepted, define $\widehat{M}^*_{1-\alpha} = M$ to be the superior set of models (the MCS).
ii) If δ_M is rejected, there is evidence that not all objects are equally good $\Rightarrow$ use the elimination rule $e_{R,M}$ to identify the object to be eliminated. Repeat steps 1 and 2 ii) until δ_M is accepted.

3 Evaluating the time-varying impact of economic data on the accuracy of stock market volatility forecasts[1]

3.1 Introduction

Forecasting volatility is a crucial part of decision-making for financial market actors as well as policy-makers. Long-horizon forecasts for volatility can be important for instance for portfolio allocation and risk management. While standard GARCH models are accurate for short-term return volatility forecasts (for example, Andersen and Bollerslev (1998)), using models which also include economic data, such as the GARCH-MIDAS model[2], have been found to be successful at longer horizons (for example, Engle et al. (2013), Conrad and Loch (2014)). There is mounting evidence that forecast accuracy varies over time (for example, Giacomini and Rossi (2010) and Stock and Watson (2003)) and that predictability varies over economic states (for example, Chauvet and Potter (2013) and for stock returns Rapach et al. (2010)).

When modelling US stock market volatility with GARCH-MIDAS models we can see that the in-sample explanatory power of economic variables varies over time.[3] In particular, the ability of many macroeconomic variables to explain stock return volatility declines over time, which motivates studying the time-variation in forecasting performance of GARCH-MIDAS models driven by economic data. The stock market volatility forecasting literature largely concentrates on evaluating the average forecasting performance over the whole out-of-sample period. Conrad and Loch (2014) considered the time-varying forecasting performance of GARCH-MIDAS models for stock market volatility and found that compared to a model driven by realised volatility macroeconomic data tends to forecast volatility well between recessions and after mid-2008. A natural question thus is how much economic data adds to forecasting per-

[1]An early version of this essay is published as *HECER Discussion Paper No. 430.*

[2]A generalised autoregressive conditional heteroscedasticity model with a mixed data sampling polynomial.

[3]See Figure 3.2 in Section 3.5.

formance at different points in time, when compared to a standard GARCH model.[4]

This paper extends the analysis in Conrad and Loch (2014) and contributes to the current literature in three ways. First, it studies the stability of the in-sample parameter estimates of GARCH-MIDAS models when estimated over rolling windows for US data, and discusses the role the weighting scheme plays in both in-sample fit and out-of-sample forecasting performance. This is crucial knowledge when estimating a model over sub-samples of the data and gives us insights into how economic data is related to volatility. In-sample results are also interesting because financial data has not been studied extensively in the GARCH-MIDAS framework.[5] Second, this paper explores the additional time-varying predictive ability provided by macroeconomic and financial variables by comparing the evolution of the out-of-sample forecasting performance of GARCH-MIDAS models to a standard GARCH model. The difference in forecasting performance thus directly reflects the impact of the economic data. In addition, to consider potential reasons for the time-variation I investigate whether the relative forecasting performance is affected by the business cycle or the market environment. Third, I consider whether this information can be used to improve the accuracy of volatility forecasts, and I determine whether forecast accuracy can be improved by combining the individual GARCH-MIDAS model forecasts. The focus of this paper is thus on improving real-time forecasts of long-term stock market volatility with the data set representing as far as possible the information set of the forecaster at the forecast origin.

I find that when forecasting over long horizons there are shifts in relative forecasting performance over time implying that (time-varying) forecast combination methods could be useful. In particular, macroeconomic variables improve predictions especially in low volatility periods but also in periods of weak economic growth. Thus the usefulness of macroeconomic variables around recessions found in earlier research, in particular when using predictive regressions, is confirmed. On the other hand, several financial variables struggle to identify a long-term component in volatility, leading to weak forecasting performance, which contrasts the important role they tend to have in predictive regressions. Financial data seems most useful for forecasting over short horizons in low volatility environments. It is clear that no single forecasting model or combination scheme excels on all horizons and in all time periods. However, as the standard GARCH model is rarely significantly better than the GARCH-MIDAS models, while, for example, forecast combinations produce forecasts which modestly, but consistently, outperform the benchmark, it seems beneficial to use economic data for long-horizon volatility forecasting. Finally, a rolling window estimation scheme leads to more volatile parameter estimates but better in-sample fit than an expanding window estimation scheme. Whether a rolling or expanding window is preferred for out-of-sample forecasting depends on the explanatory data, forecasting horizon and loss function.

The paper is organised as follows. Section 3.2 discusses the relevant literature, while Section 3.3 presents the GARCH-MIDAS model and the forecasting set-up. The

[4]It is well established, for example in Lindblad (2017), that realised volatility produces good in-sample fit but performs badly out-of-sample compared to a standard GARCH model.

[5]While Asgharian et al. (2013) included financial data and found some of it useful for forecasting, they concentrated on reporting results for the principal components.

data set is introduced in Section 3.4, and Section 3.5 briefly establishes in-sample results. When discussing the out-of-sample results in Section 3.6 I first present baseline full-sample results, before considering the time-variation in forecasting performance. Section 3.7 considers whether conditioning on current economic circumstances improves forecasts. I evaluate forecast combination methods in Section 3.8 before concluding in Section 3.9.

3.2 Volatility forecasting using economic predictors

When forecasting stock return volatility focus has been on one-period-ahead forecasts where the step tends to be relatively short. For example, Engle (1982), Bollerslev (1986), Andersen and Bollerslev (1998) and Hansen and Lunde (2005) all considered one-step-ahead forecasts. Over short horizons GARCH(1,1) models usually perform well. Poon and Granger (2003) thoroughly reviewed the volatility forecasting literature, and I will here concentrate on the literature considering long-horizon forecasts and models incorporating economic data. Ghysels et al. (2009) discussed multi-horizon volatility forecasts, comparing iterated, direct and MIxed DAta Sampling (MIDAS) approaches to the commonly used rule-of-thumb, where volatility is scaled up by the number of trading days. They found that for long horizons (over 30 days ahead) the MIDAS regression forecasts dominate. Ghysels et al. (2009) thus argued that volatility is forecastable also over long horizons contrary to the evidence in Christoffersen and Diebold (2000). Their study does not, however, consider GARCH-MIDAS models or include macro-finance variables to enhance volatility forecasts.

There is ample evidence that stock return volatility is higher in recessions than in expansions (for example, Schwert (1989)). Nevertheless, mixed results on the usefulness of economic data for modelling and forecasting volatility is found in, for example, Davis and Kutan (2003), Errunza and Hogan (1998), Pierdzioch et al. (2008) and Paye (2012). Other papers, such as Hamilton and Lin (1996), Cakmakli and van Dijk (2010), Christiansen et al. (2012) and Diebold and Yilmaz (2008) were more successful in linking economic developments to return volatility. These papers mostly rely on predictive regressions and VARs. Papers building on the component GARCH framework, introduced by Engle and Lee (1999), have also successfully linked macroeconomic variables and stock market volatility. In particular, Engle et al. (2013) introduced the GARCH-MIDAS model, which decomposes volatility into a short-term component that fluctuates around a long-term trend determined by economic data. They compared forecast accuracy of GARCH-MIDAS models driven by industrial production and the producer price index to a GARCH-MIDAS model driven by realised volatility over several (long) subsamples, using full-sample parameter estimates. Conrad and Loch (2014), Asgharian et al. (2013), Asgharian et al. (2015) and Lindblad (2017) used the GARCH-MIDAS model to show that economic data helps explain and forecast stock return volatility.

Following the literature on time-variation in the accuracy of macroeconomic (for example, Stock and Watson (2003)) and stock return (for example, Rapach et al. (2010)) forecasts, it is natural to think that the ability of economic data to forecast return volatility could be time-varying and depend on the state of the business cycle or the volatility environment. Christiansen et al. (2012), compared the dynamic out-of-sample performance of predictive regressions, estimated using a Bayesian Model Averaging approach, to autoregressive benchmarks, concluding that macro-finance variables add to predictability over the most recent financial crisis period. Paye (2012) found using predictive regressions that macroeconomic variables are especially useful for forecasting volatility around recessions. The closest previous contribution to this paper is Conrad and Loch (2014), who used the GARCH-MIDAS model and considered the (unconditional) predictive ability test by Giacomini and White (2006) for volatility forecasts calculated over a two year rolling window for three different daily horizons. In contrast, I use the Fluctuation test methodology presented in Giacomini and Rossi (2010), which is a rolling window version of the test by Giacomini and White (2006), advocating a relatively large window size, on monthly volatility forecasts. My choice of evaluation window relies on the considerations of good size and power properties of the Fluctuation test discussed in Giacomini and Rossi (2010). Another significant difference to the time-varying forecast comparison is the benchmark model, which in Conrad and Loch (2014) was a GARCH-MIDAS model driven by realised volatility (RV), while I use a standard GARCH model. This is a much tougher benchmark to beat (see, for example, Lindblad (2017)), and it also allows us to directly infer the marginal benefit of economic data for forecasting volatility at each point in time. Conrad and Loch (2014) found that models using macroeconomic data lead to significantly more accurate forecasts than the GARCH-MIDAS-RV model between the past two recessions and since the beginning of the financial crisis.

3.3 Methodology

3.3.1 The GARCH-MIDAS model

The GARCH-MIDAS model by Engle et al. (2013) is a multiplicative two-component model for the conditional variance, where the high-frequency component is modelled as a standard GARCH process while the low-frequency component is determined by economic data. The high-frequency component can be thought of as fluctuating around a slow-moving long-term trend driven by variables evolving at a lower frequency than returns. The MIxed DAta Sampling (MIDAS) approach, introduced by Ghysels et al. (2004)[6], deals with the challenges related to using data sampled at different frequencies within the same model. A key feature of MIDAS is capturing the lag structure of the explanatory data by a known function which depends on only a few parameters.

Following the interpretation in Engle and Rangel (2008), which builds on the log-linear dividend-ratio model in Campbell (1991) and Campbell and Shiller (1988), the stock return on day i and in period (month or quarter) t can be modelled as having a multiplicative specification for the conditional variance:

$$r_{i,t} = E_{i-1,t}(r_{i,t}) + \sqrt{\tau_{i,t}\, g_{i,t}}\, \varepsilon_{i,t}, \quad \varepsilon_{i,t} \mid \Phi_{i-1,t} \sim N(0,1), \quad \forall i = 1, ..., N_t$$

where $\Phi_{i-1,t}$ represents the information set up to day $i-1$, and N_t is the number of

[6] Discussed in detail in Ghysels et al. (2004), Ghysels et al. (2005), Ghysels et al. (2006), Ghysels et al. (2007), Andreou et al. (2010), and Wang and Ghysels (2015).

trading days in period t. $\sigma^2_{i,t} = \tau_{i,t}\ g_{i,t}$ is the total conditional variance, where τ_t[7] is the long-term volatility component and $g_{i,t}$ the high-frequency (GARCH) component. Following Engle et al. (2013) the expected return is assumed constant: $E_{i-1,t}(r_{i,t}) = \mu$.

It is well established that stock return volatility is asymmetric (see, for example, Awartani and Corradi (2005) and the references therein), i.e., that positive and negative news have different impact on volatility. Therefore I use the asymmetric GJR-GARCH model (by Glosten et al. (1993)):

$$g_{i,t} = \omega + (\alpha + \gamma D_{i-1,t})\frac{(r_{i-1,t} - \mu)^2}{\tau_t} + \beta g_{i-1,t} \tag{3.1}$$

where $D_{i-1,t}$ is an indicator function, taking the value 1 when $(r_{i-1,t} - \mu) < 0$ and 0 otherwise. Thus, γ describes the degree of asymmetry in volatility. ω is normalised to $\omega = 1 - \alpha - \beta - \gamma/2$ so that $E(g_{i,t}) = 1$. To ensure stationarity the condition $\alpha + \beta + \gamma/2 < 1$ is imposed. In addition, $\alpha > 0$, $\beta \geq 0$ and $\alpha + \gamma \geq 0$ is assumed to ensure the variance remains positive.

Following Engle et al. (2013) the MIDAS polynomial with one explanatory variable (X, which is, for example, macroeconomic data) takes the form:

$$log\ \tau_t = m + \theta \sum_{k=1}^{K} \varphi_k(\omega_1, \omega_2) X_{t-k} \tag{3.2}$$

where $\varphi_k(\omega_1, \omega_2)$ is a weighting scheme and K is the number of lags of explanatory data included. The logarithmic specification ensures non-negativity of the long-term volatility component (τ_t) even when the explanatory variable takes negative values. If the variable does not affect stock market volatility (i.e., $\theta = 0$), all volatility is captured by the short-term component and the model collapses to the GJR-GARCH model with $\tau_t = m$. The standard GARCH model is therefore nested in the GARCH-MIDAS-X specification. The sign of θ is interpretable: $\theta > 0$ ($\theta < 0$) implies that higher values of X are linked to higher (lower) long-term volatility in stock returns.

A commonly used flexible but parsimonious weighting scheme is the beta lag polynomial[8], which guarantees positive weights (thus ensuring non-negativity of volatility) that add up to one (this normalisation allows identifying θ):

$$\varphi_k(\omega_1, \omega_2) = \frac{(\frac{k}{K})^{\omega_1 - 1}(1-\frac{k}{K})^{\omega_2 - 1}}{\sum_{j=1}^{K}(\frac{j}{K})^{\omega_1 - 1}(1-\frac{j}{K})^{\omega_2 - 1}}, \quad \text{where } \sum_{k=1}^{K} \varphi_k(\omega_1, \omega_2) = 1.$$

The weight parameters, ω_1 and ω_2, govern the shape of the weighting scheme and can be estimated or fixed before estimation. The beta polynomial allows both monotonously decreasing weights ($\omega_1 = 1$) and hump-shaped weights ($\omega_1 < \omega_2$). If $\omega_1 = 1$, the rate of decay is determined by ω_2. A larger value of ω_2 indicates faster decay. If $\omega_2 < \omega_1$, all weight can be on distant lags, which can be seen as counterintuitive. If $\omega_1 = \omega_2 = 1$, the weights are equal ($1/K$) for all lags, which corresponds to a moving average.

[7] $\tau_{i,t}$ is fixed for all i in period t. Hence, the subscript i is suppressed to ease notation and emphasise that τ_t evolves at a lower frequency than $g_{i,t}$.

[8] Weighting schemes are discussed in more detail in Ghysels et al. (2007).

To assess how much the variation in a particular variable explains of the overall expected volatility, Engle et al. (2013) suggested calculating variance ratios: $\frac{Var(log(\tau_t))}{Var(log(\tau_t g_{i,t}))}$. The variance ratio can be interpreted as a measure of fit in the sense that the higher the variance ratio is, the larger is the share of the total expected volatility that can be explained by the variation in the long-term component. The GARCH-MIDAS model can be estimated using maximum likelihood or quasi-maximum likelihood if the assumption of normally distributed errors does not hold.[9]

3.3.2 Forecasting with the GARCH-MIDAS model

The one-step ahead volatility prediction is given directly by equations 3.1 and 3.2. For further horizons I iterate forward the daily GJR-GARCH model forecasts and combine this short-term forecast with a forecast for the long-term component (τ_t). For the GJR-GARCH model the forecast for day i is formed as:

$$E\left[g_{i,t}|F_{N_{t-1},t-1}\right] = 1 + (\alpha + \beta + \gamma/2)^{i-1}(g_{1,t} - 1), \tag{3.3}$$

where N_t is the number of trading days in period t, and $F_{N_{t-1},t-1}$ denotes the information set in period $t-1$. The forecast for total volatility for period t can be expressed as:

$$E\left[\sum_{i=1}^{N_t} g_{i,t}\tau_t\varepsilon_{i,t}^2|F_{N_{t-1},t-1}\right] = \tau_t\left[N_t + (g_{1,t} - 1)\frac{1-(\alpha+\beta+\gamma/2)^{N_t}}{1-\alpha-\beta-\gamma/2}\right]. \tag{3.4}$$

Following Conrad and Loch (2014) I create non-overlapping monthly forecasts by summing the daily forecasts over the respective month while keeping τ_t fixed at its one-step ahead prediction for all horizons. Because the forecast of the GARCH component converges to its (constant) unconditional expectation as the forecast horizon increases, in the long run the forecast differences are entirely driven by the long-term components.

3.3.3 Forecasting set-up

The GARCH-MIDAS model has relatively many parameters to estimate, and therefore the estimation period needs to be long enough. However, in order to detect time-variation in the out-of-sample forecasts the evaluation period needs to be as long as possible. I thus divide the whole sample (January 1973 - June 2017) roughly into half: the first estimation period is January 1973 - December 1994, and the out-of-sample evaluation period is January 1996 - June 2017. As the short-term GARCH components are similar across all GARCH-MIDAS specifications, the largest gains in forecasting from including economic variables is expected to be achieved over long horizons. I therefore consider forecast horizons from 1 month up to 12 months.

For the out-of-sample evaluation I use a rolling window estimation scheme, i.e., the estimation window is shifted forward by one period and the model is re-estimated

[9]While consistency and asymptotic normality of the QML estimator for the rolling window GARCH-MIDAS model with realised volatility was established in Wang and Ghysels (2015), it has not been shown for the more general GARCH-MIDAS model with macroeconomic variables.

before the next set of forecasts is calculated. A rolling window estimation scheme takes into account potential parameter instability, which is important if the relationship between long-term stock market volatility and the economic variables changes over time. More importantly, as the forecast comparison methods used require limited memory estimators, an expanding window is not feasible.

The forecasts are evaluated against realised volatility calculated as the monthly sum of squared daily returns ($RV_t = \sum_{i=1}^{N_t} r_{i,t}^2$). Forecast accuracy of a model is measured as the squared forecast error as well as the absolute value of the forecast error. Squared forecast errors put significant weight on the largest forecast errors, which is useful if one wants to emphasise large forecast errors over smaller ones. In addition, Patton (2011) argues that while the mean squared forecast error (MSFE) loss function is robust in the sense that using a noisy proxy for volatility, such as the sum of squared daily returns, does not change the ranking of forecasting models, the mean absolute forecast error (MAFE) loss function is not. However, Poon and Granger (2003) note that when using squared returns as the quantity of interest and using squared errors as the measure of forecast accuracy, one is effectively comparing the fourth moments of the data, which can complicate the comparison. In general, the results are similar, although as expected statistical significance is weaker for squared forecast errors.

A natural benchmark is the GJR-GARCH(1,1) model since it is nested in the GARCH-MIDAS specification. The GJR-GARCH(1,1) benchmark thus reveals whether economic variables, including realised volatility, are useful for forecasting stock return volatility. It is therefore a more natural benchmark than the GARCH-MIDAS model driven by realised volatility when studying the time-variation in the impact economic data has on the accuracy of stock market volatility forecasts. In addition, previous research[10] has found that the GJR-GARCH(1,1) model tends to outperform the GARCH-MIDAS model driven by realised volatility when forecasting long-term volatility.

3.3.4 Evaluating the time-variation in relative forecasting performance

The accuracy of the forecasting framework is important, but there is often considerable uncertainty regarding the choice of model. It is, therefore, important to be able to test the relative forecasting performance of competing models and to this end several frameworks have been developed.[11] However, as pointed out by Giacomini and Rossi (2010), the relative forecasting performance of models might be time-varying due to, for example, structural instability. Whether the relative forecasting performance of two models has shifted over time is an interesting and important question to complement full-sample results. To this end Giacomini and Rossi (2010) proposed the Fluctuation test, where the idea is to compare scaled and centred h-step-ahead out-of-sample forecast losses calculated over rolling windows of size m:

$$F_{t,m} = \hat{\sigma}^{-1} m^{1/2} \sum_{j=t-m/2}^{t+m/2-1} \Delta L_j(\hat{a}_{1,j-h,R}, \hat{a}_{2,j-h,R}), \tag{3.5}$$

[10] See, for example, Conrad and Loch (2014) or Lindblad (2017).

[11] For example, Diebold and Mariano (1995), West (1996), McCracken (2000), Clark and McCracken (2001), Clark and West (2006) and Giacomini and White (2006).

where $t = R + h + m/2, \ldots, T - m/2 + 1$, R is the in-sample size, ΔL_j is the difference in two loss functions in period j, $\hat{\sigma}^2$ is a HAC estimator of the variance (σ^2) and $\hat{a}_1$ and $\hat{a}_2$ are the in-sample parameter estimates of each model.[12] The Fluctuation test tests the null hypothesis that the local relative forecasting performance equals zero at each point in time: $H_0 : E[\Delta L_t(\hat{a}_{1,t-h,R}, \hat{a}_{2,t-h,R})] = 0$ and is therefore a rolling window version of the Giacomini and White (2006) test. The testing framework allows both nested and non-nested models as well as non-linear models, but the parameters need to be estimated using a limited memory estimation scheme, such as rolling windows. Giacomini and Rossi (2010) showed that if the ratio between m and $T - R$ (out-of-sample size) is too small the Fluctuation test is oversized. The size and power properties of the test are found to be good for $\frac{m}{T-R} \approx 0.3$. As my out-of-sample size is 258 I need, for example, $m = 78$, which corresponds to 6.5 years of monthly data. The test is therefore designed to detect long-term shifts in forecasting performance.

3.4 Data

I use the continuously compounded daily stock market return on the CRSP index from January 1973 to June 2017. From a theoretical perspective time-variation in stock return volatility can be linked to uncertainty regarding future cash flows, which can stem from, for example, uncertainty regarding the true macroeconomic situation and expectations regarding the future economic environment. As explanatory variables I include a representative set of commonly used (monthly) predictors for stock return volatility, covering the financial markets, the macroeconomy and expectations regarding the economic environment. While the important role of many macroeconomic variables in driving long-term volatility has been well established in the GARCH-MIDAS literature (see Section 3.2), financial variables, with the exception of the term spread and realised volatility, have received less attention. The exception is Asgharian et al. (2013), who included the 3-month T-bill rate, a default spread and the exchange rate as well, and found that the interest rate and the default spread seemed useful for forecasting. The results, however, concentrated on principal components based on macroeconomic and financial data. Using predictive regressions financial variables have been identified as important predictors of stock return volatility (for example, Christiansen et al. (2012)).

The macroeconomic variables included are real-time housing starts (change in level), the real-time Aruoba-Diebold-Scotti Business Conditions index (ADS index)[13], the Buying Conditions index (forward-looking sub-index of the University of Michigan consumer confidence index, change in level), and the ISM New Orders index (level). As a forward-looking indicator housing starts has been among the best predictors for stock return volatility (for example, Conrad and Loch (2014)), the ADS index reflects the current economic situation, and the Buying Conditions index and the ISM New Orders index represent expectations of the macroeconomic situation by different sectors of the economy.

[12]See Giacomini and Rossi (2010) for details.

[13]Includes, for example, industrial production and labour market data. Prior to 2008 real-time vintages are unavailable.

As financial data I include commonly used equity return predictors (as in, for example, Goyal and Welch (2008)) and in particular those found most useful for predicting stock return volatility in, for example, Christiansen et al. (2012) and Conrad and Loch (2014).[14] I use a realised volatility measure (sum of the absolute value of daily returns: $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$), which performed well in-sample in Lindblad (2017), and the term spread (difference between the 10-year Treasury bond yield and the 3-month T-bill rate), which was found to be an important driver of long-term volatility already in Conrad and Loch (2014). In addition, I include the short-term and long-term interest rates (level and change over month) as well as the default spread (default risk of corporate bonds measured as the difference between BAA and AAA bond yields) as a measure of credit risk. The short-term interest rate and the default spread were found useful for forecasting volatility in Asgharian et al. (2013). To capture equity market movements I include excess market returns. For missing values I use the previous month's data.[15] See Appendix 3.A for all data sources.

To determine whether a broad set of macroeconomic and financial variables is useful for forecasting stock market volatility I use the dataset and methodology in McCracken and Ng (2016) to extract factors using principal components analysis. The dataset currently comprises 128 macroeconomic and financial variables. I use the first four principal components (PC), which explain a combined 34% of the total variation in the data, in the analysis.[16] As shown in detail in Appendix 3.B, the first PC relates to real activity and employment, the second one concentrates on price variables, the third one relates mainly to interest rate spreads, while the fourth one is dominated by financial variables. I use as far as possible real-time data for the principal components in the rolling window analysis. Historical vintages go back to August 1999. Before that I use the August 1999 vintage and recursively estimate the PCs for each period, using only historical data. The first time-varying PC relates mainly to the same underlying macroeconomic series – real activity and employment related series – as the full-sample PC, as shown in Figure 3.13. For the second and third PCs the compositions vary, although the interpretation of the factors remains relatively constant over time. The second PC mainly relates to interest rates and interest rate spreads but also to price variables. For the third PC one cluster relates to price variables, a second one to interest rates and a third one relates to housing market data.

3.5 In-sample results

First, I establish in-sample results for the full sample period. Then, I look at parameter stability over the forecasting horizon using a rolling window estimation scheme.

[14] A requirement is that data is available from January 1971 until June 2017 (up to two years of economic data is needed to estimate the model for the first period). Therefore, for example, the investor sentiment index by Baker and Wurgler (2006) (available until September 2015) and the E/P and D/P ratios are not included, although they have been successful predictors of returns. In results which are available upon request I determine that these variables are not important drivers of long-term stock market volatility over the sample period for which they are available.

[15] This is important for the Buying Conditions index, available only at a quarterly frequency before 1978.

[16] See McCracken and Ng (2016) for details on the data, the extracted factors (which are very similar to those extracted here) and the methodology.

Importantly, it will reveal how the long-term relationship between economic variables and stock market volatility has changed over time, as identified by the GARCH-MIDAS model.

3.5.1 Full-sample results

In the MIDAS polynomial lag length K needs to be determined. I choose between K = 12 and K = 24 for each model, i.e., one or two years of lagged economic data and proceed with the lag length maximising the log-likelihood function value.[17] The same K is used throughout the estimations.

Table 3.1 presents the in-sample estimation results over the full sample of all the GARCH-MIDAS models and the baseline GJR-GARCH model. The macroeconomic data, the term spread and realised volatility get highly significant estimates for θ as well as high variance ratios, implying the variables are useful for modelling stock market volatility.[18] These results largely echo earlier results in Conrad and Loch (2014) and Lindblad (2017).

The interest rate data does not lead to good in-sample fit, as evidenced by both weakly significant parameter estimates and low variance ratios, and as such these models are unlikely to produce forecasts very different from the baseline GJR-GARCH(1,1) model. The consequence of this is displayed in Figure 3.1, where the long-term volatility component of the GARCH-MIDAS models driven by the 3M T-bill rate is basically a horizontal line. Therefore, only the 3M T-bill rate (level), which has the highest variance ratio and a highly significant parameter estimate, is included in the subsequent out-of-sample analysis.[19]

The default spread, 3M T-bill rate and excess market return get positive estimates for θ, implying that a higher risk of default, a higher interest rate and a higher excess market return lead to higher stock return volatility. The first principal component explains a large, 16% share of the total variance, while the two following PCs explain roughly 10% each. The estimates for θ are also highly significant. On the other hand, the fourth PC has a clearly lower variance ratio, implying it is not an important driver of long-term volatility. This is in line with it being mainly related to financial data. I thus proceed using the first three PCs.

Figure 3.2 shows how the in-sample explanatory power of various GARCH-MIDAS models varies over time, as indicated by the variance ratio calculated over rolling windows. The GARCH-MIDAS model where the long-term volatility component is driven by lagged realised volatility (RV) explains a stable 40%–50% of total volatility, while the long-term component of the model driven by the term spread explains a relatively

[17] The results are not, however, materially changed by the choice of 12 or 24 lags.

[18] When testing the significance of θ, θ and the weight parameters ω_i are not separately identified under the null hypothesis, which affects the asymptotic distribution of the test statistic. However, I follow the convention in the GARCH-MIDAS literature (for example, Engle et al. (2013) and Conrad and Loch (2014)) and proceed using the standard t-statistic. In addition, Appendix 3.F discusses estimates of θ using a fixed weighting scheme. See Ghysels et al. (2007) for a discussion of the problem in MIDAS regressions.

[19] The weighting scheme of the 3M T-bill rate (level) can be seen as counterintuitive, putting all weight on a distant lag and with the parameter estimate for ω_1 reaching the upper bound of the parameter space. However, I am mostly interested in the rolling window estimates of the model parameters, discussed in Section 3.5.2.

Table 3.1: Estimation results of the GARCH-MIDAS-X model

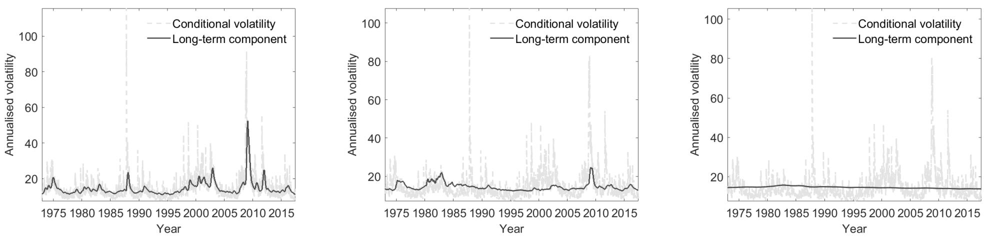
	μ	α	β	γ	θ	ω_1	ω_2	m	VR	LLF
GJR	0.0466***	0.0217***	0.9024***	0.1073***	-	-	-	0.8500***	-	-14281.88
	(0.0074)	(0.0050)	(0.0136)	(0.0180)				(0.0872)		
RV	0.0482***	0.0133**	0.8559***	0.1438***	0.0639***	2.5509***	6.5085**	-1.1786***	34.79	-14229.19
	(0.0073)	(0.0056)	(0.0175)	(0.0203)	(0.0050)	(0.9418)	(2.6167)	(0.0979)		[0.0094]
BC	0.0454***	0.0182***	0.8936***	0.1174***	-0.1788***	1.8624***	2.1397***	-0.1588*	14.24	-14253.96
	(0.0074)	(0.0052)	(0.0143)	(0.0186)	(0.0259)	(0.4391)	(0.7643)	(0.0844)		[0.0010]
ISM	0.0456***	0.0144***	0.8987***	0.1188***	-0.0522***	1	2.6036***	2.6681***	15.37	-14254.51
	(0.0074)	(0.0054)	(0.0138)	(0.0183)	(0.0086)		(0.9094)	(0.4760)		[0.6112]
ADS	0.0464***	0.0159***	0.8968***	0.1174***	-0.4817***	1	3.3587***	-0.2496***	15.09	-14255.13
	(0.0074)	(0.0053)	(0.0138)	(0.0183)	(0.0761)		(0.8423)	(0.0874)		[0.6067]
HS	0.0463***	0.0170***	0.8952***	0.1179***	-0.0150***	2.0944***	1.7774***	-0.2137**	17.59	-14249.16
	(0.0074)	(0.0052)	(0.0143)	(0.0185)	(0.0022)	(0.6512)	(0.4682)	(0.0851)		[0.0000]
TS	0.0468***	0.0174***	0.8933***	0.1174***	-0.2485***	2.8814	1.6183*	0.2411**	13.87	-14255.26
	(0.0073)	(0.0052)	(0.0149)	(0.0192)	(0.0417)	(2.5458)	(0.8912)	(0.1095)		[0.0148]
DS	0.0456***	0.0133***	0.8977***	0.1217***	0.5605***	1	6.7455**	-0.8116***	12.07	-14261.93
	(0.0073)	(0.0051)	(0.0144)	(0.0193)	(0.0994)		(2.9512)	(0.1500)		[0.2594]
3MTb	0.0456***	0.0177***	0.9028***	0.1127***	0.0437***	300	233.5683	-0.3906***	4.46	-14273.72
	(0.0073)	(0.0052)	(0.0139)	(0.0187)	(0.0157)	(499.2185)	(402.5675)	(0.1278)		[0.0441]
3M Δ	0.0458***	0.0175***	0.9020***	0.1126***	-0.7768**	1	1.7220*	-0.1821*	3.18	-14275.57
	(0.0074)	(0.0052)	(0.0135)	(0.0181)	(0.3249)		(0.8999)	(0.0959)		[0.0795]
10Y	0.0462***	0.0203***	0.9030***	0.1090***	0.0221	1	1.0000	-0.3145**	0.83	-14280.65
	(0.0073)	(0.0051)	(0.0137)	(0.0183)	(0.0185)		(3.7585)	(0.1601)		[0.3694]
10Y Δ	0.0467***	0.0204***	0.9029***	0.1090***	-0.6228	5.2828*	34.9021	-0.1704*	2.20	-14275.41
	(0.0074)	(0.0050)	(0.0132)	(0.0176)	(0.3525)	(2.5327)	(22.6232)	(0.1016)		[0.0032]
EMR	0.0479***	0.0159***	0.9066***	0.1165***	0.1089***	1	3.8440***	-0.2337**	9.30	-14262.94
	(0.0073)	(0.0048)	(0.0112)	(0.0157)	(0.0286)		(0.8528)	(0.1135)		[1.0000]
PC1	0.0466***	0.0163***	0.8944***	0.1194***	0.9380***	1	6.9868**	-0.2252***	16.17	-14254.73
	(0.0074)	(0.0053)	(0.0141)	(0.0185)	(0.1450)		(2.9702)	(0.0847)		[0.6032]
PC2	0.0459***	0.0174***	0.8970***	0.1171***	-1.8320***	12.0342	6.2691	-0.1859**	10.01	-14263.10
	(0.0074)	(0.0053)	(0.0144)	(0.0194)	(0.4827)	(28.1765)	(16.0309)	(0.0902)		[0.0216]
PC3	0.0458***	0.0172***	0.8988***	0.1154***	1.0902***	4.5780	2.3205	-0.1804**	10.55	-14263.59
	(0.0074)	(0.0052)	(0.0145)	(0.0191)	(0.2500)	(3.7068)	(1.6009)	(0.0913)		[0.0031]
PC4	0.0467***	0.0210***	0.9012***	0.1089***	-0.7049**	12.4661	30.6975	-0.1723*	2.47	-14276.94
	(0.0074)	(0.0051)	(0.0137)	(0.0181)	(0.3452)	(8.3094)	(26.2808)	(0.0987)		[0.0027]

*Bollerslev-Wooldridge QMLE robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. VR is the variance ratio from Section 3.3.1 multiplied by 100. MIDAS polynomial:* $\log \tau_t = m + \theta \sum_{k=1}^{K} \varphi_k(\omega_1, \omega_2) X_{t-k}$. *All models are estimated with a restricted* ($\omega_1 = 1$) *and an unrestricted weighting scheme. The model reported in the table is chosen based on a likelihood ratio test between the restricted and unrestricted specifications. The related p-value is reported below the value of the log likelihood function (LLF). The used variables with their abbreviations are: GJR-GARCH(1,1) (GJR), Realised volatility (RV), Buying Conditions index (BC), ISM New Orders index (ISM), ADS index (ADS), Housing starts (HS), Term spread (TS), Default spread (DS), 3M T-bill rate, level (3MTb), 3M T-bill rate, change over month (3M Δ), 10 year Treasury yield, level (10Y), 10 year Treasury yield, change over month (10Y Δ), Excess market return (EMR), Principal component (PC).* $K = 24$ *for all models except those including RV, DS, 3M Δ and EMR, for which* $K = 12$.

stable 20%–30%. For the remaining models the explanatory power of the economic variables seems to decline over time.

3.5.2 Parameter instability

There are 270 out-of-sample months (January 1995 - June 2017) and hence 270 estimates for each parameter. In this section I discuss how the parameter estimates vary over time and how representative the full-sample results are. I examine whether the choice of restricted or unrestricted weighting scheme remains constant over the out-of-sample



(a) Realised volatility (b) Default spread (c) 3M T-bill rate

Figure 3.1: Total and long-term volatility components for a selection of models from Table 3.1. Similar graphs for the remaining models are available upon request.

period. Taking into account parameter instability can be important for forecasting. Parameter instability could signal structural breaks, which indicates that the relationship between stock return volatility and the economic variables changes over time.

Regarding the choice between a restricted and unrestricted weighting scheme, Table 3.2 presents the percentage of times the unrestricted weighting scheme is chosen over the restricted one, determined by a likelihood ratio test in each of the 270 out-of-sample periods.[20] Clearly, for realised volatility, the ISM New Orders index and PC1 the restricted model is always chosen, while for the Buying Conditions index and housing starts we always choose the unrestricted weighting scheme. For the ADS index, the default spread and the 3M T-bill rate we almost always choose the restricted weighting scheme. On the other hand, for the term spread, the excess market return, PC2 and PC3 the choice varies, although the unrestricted weighting scheme is chosen more often.[21] The chosen weighting schemes mostly confirm the full-sample results in Section 3.5, except for realised volatility and the 3M T-bill rate, for which a restricted weighting scheme now gets more support, and excess returns, for which we would now mostly choose an unrestricted weighting scheme. For realised volatility the restricted weighting scheme is in line with the earlier literature, and for the 3M T-bill rate it was noted already earlier that the optimal unrestricted weighting scheme led to counterintuitive weights.

Figure 3.3 illustrates the time-variation in the estimated GARCH parameters as well as the time-variation in the statistical significance of γ, which describes the degree of asymmetry in returns. The parameters relating to the GARCH model behave very similarly over time and in line with the GJR-GARCH(1,1) model. The exception is the GARCH-MIDAS-RV model, for which especially β is estimated lower and γ higher compared to the other models. Interestingly, γ roughly doubles in magnitude over time in all models. This implies that smaller-than-expected returns (with estimated

[20]The graphs in this section as well as the out-of-sample analysis are based on the weighting scheme which is chosen more often. Appendix 3.C discusses in more detail the time-variation in and the implications of the choice of weighting scheme. Overall, the differences are mostly relatively small and depend on the variable, forecasting horizon and loss function.

[21]Notice that the variation in the optimal weighting scheme for the principal components can also be a result of the changing composition of the time-varying PC itself.

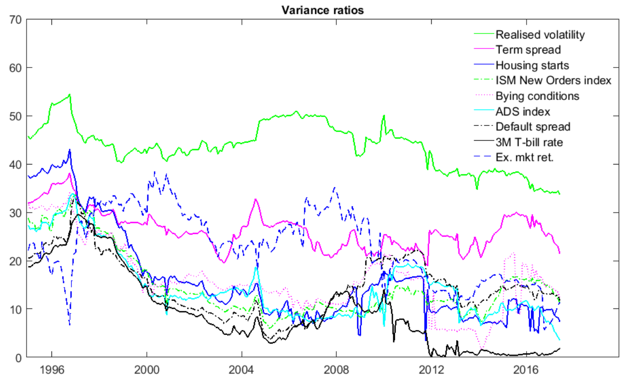
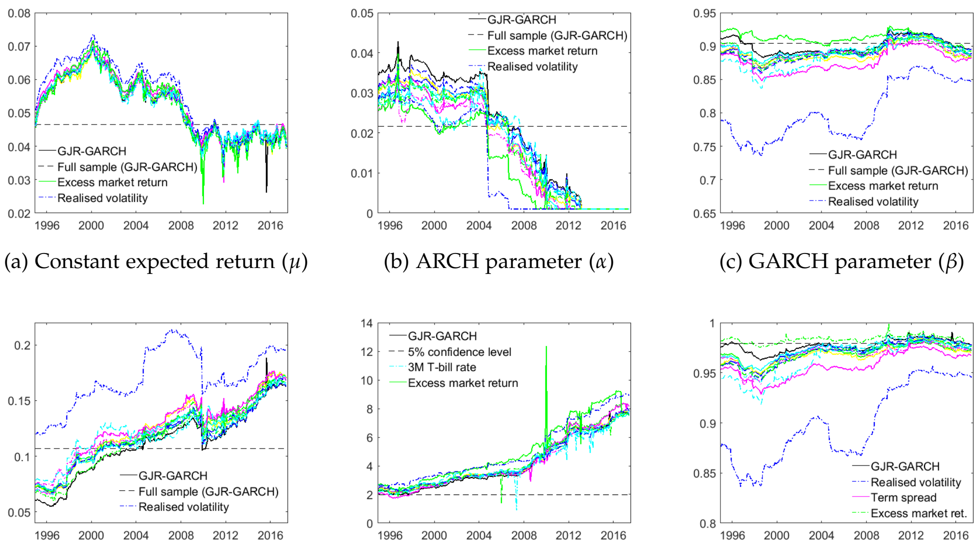


Figure 3.2: Variance ratios of selected GARCH-MIDAS models. Based on monthly rolling windows: first period January 1973 - December 1994, last period July 1994 - June 2016.



(a) Constant expected return (μ)

(b) ARCH parameter (α)

(c) GARCH parameter (β)

(d) Asymmetry parameter (γ)

(e) t-statistics for γ

(f) $\alpha + \beta + 0.5\gamma$

Figure 3.3: Time-variation in GJR-GARCH model parameter estimates. Legends contain selected series.

parameter $\alpha + \gamma$) affect volatility much more than larger-than-expected returns (with estimated parameter α), and this effect becomes more pronounced towards the end of the sample period. γ also remains significantly different from zero for all models in most periods (see Panel 3.3e).

The relationship between economic data and stock return volatility is described by

Table 3.2: Choice between restricted and unrestricted weighting scheme

	% of total		% of total		% of total
Buying Conditions index	100	ADS index	4.81	Realised volatility	0
Housing starts	100	3M T-bill rate	7.41	Excess market return	69.63
ISM New Orders index	0	Default spread	2.22	Term spread	54.44
PC 1	0	PC 2	65.56	PC 3	78.89

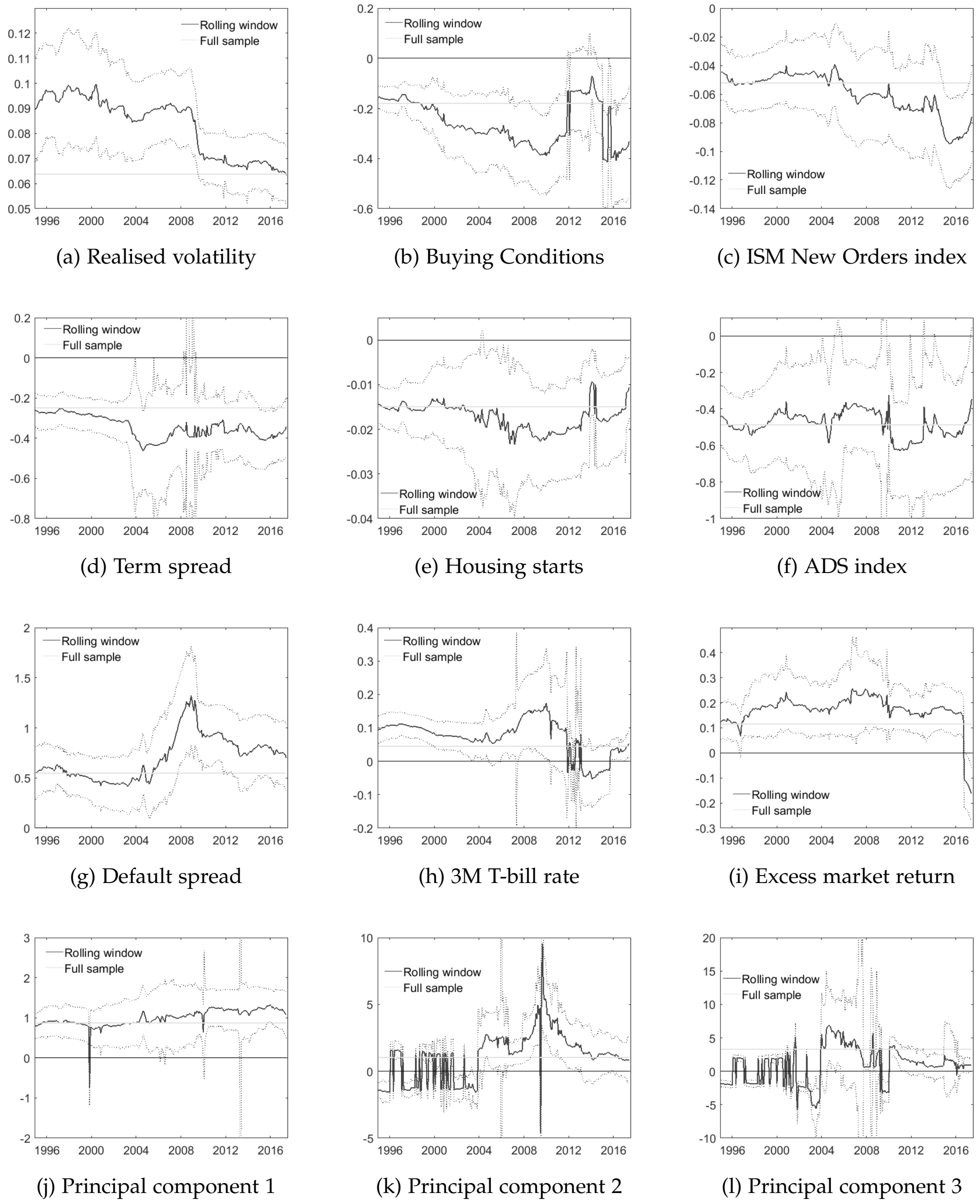
The table reports the percentage of times the unrestricted weighting scheme is chosen over the restricted one, i.e., if the number is over 50 the unrestricted weighting scheme is chosen more often. Choice is based on a likelihood ratio test (5% confidence level) in each of the 270 out-of-sample periods.

θ. Figure 3.4 shows how the estimates for θ change over the out-of-sample period in the different GARCH-MIDAS specifications. Mostly θ fluctuates around the full-sample estimate, but, in particular for realised volatility, there is a time trend in θ, indicating a rolling estimation scheme is appropriate. Counterintuitively the sign of θ for the excess market return changes at the end of the sample period. For the second and third principal components the sign of θ varies over time, resulting, most likely, from the time-varying correlation with the underlying economic variables.[22] In many specifications θ is significantly different from zero in most periods, confirming that economic data is important for long-term volatility. The main exceptions are the second and third principal components and the 3M T-bill rate, for which θ is, especially recently, not significantly different from zero at the 5% level, the basic GJR-GARCH model could be used instead.

It is also interesting to consider how the weight parameter(s) in the GARCH-MIDAS specifications change over time. Figure 3.5 depicts the time-variation in the estimated weight parameters for each of the GARCH-MIDAS models. The weight parameter (ω_2) for realised volatility and the ISM New Orders index is shrinking, implying that the decay of the weights becomes slower and further lags become increasingly important. The ADS index, the 3M T-bill rate, the default spread and the first PC mostly exhibit a similar weighting scheme to the full-sample results, but there are time periods when only the most recent data matters (i.e., ω_2 is very large). For the term spread and the Buying Conditions index, and to a lesser degree housing starts, towards the end of the sample period there is a tendency for the weighting schemes to put significant weight on a specific lag, which is not necessarily the most recent one.

It is clear that most of the economic data are important for volatility in most subperiods. The time-variation in both the weighting schemes and the estimates for θ indicate that the relationship between economic data and long-term stock market volatility varies over time and that the chosen sample period matters. The strong variation in weights over time can reflect estimation problems (related to, for instance, the relatively small estimation window size) but can also be due to a changing relationship between the variables and volatility. This is of particular concern for the GARCH-MIDAS models driven by the excess market return, the term spread and the third PC, for which several of the weight parameters are imprecisely estimated and hit the upper bound (300) used in the estimation. To guard against estimation problems impacting

[22] Appendix 3.F shows that the time-variation in the sign of the parameter estimates is not a result of the chosen weighting scheme.



(a) Realised volatility (b) Buying Conditions (c) ISM New Orders index

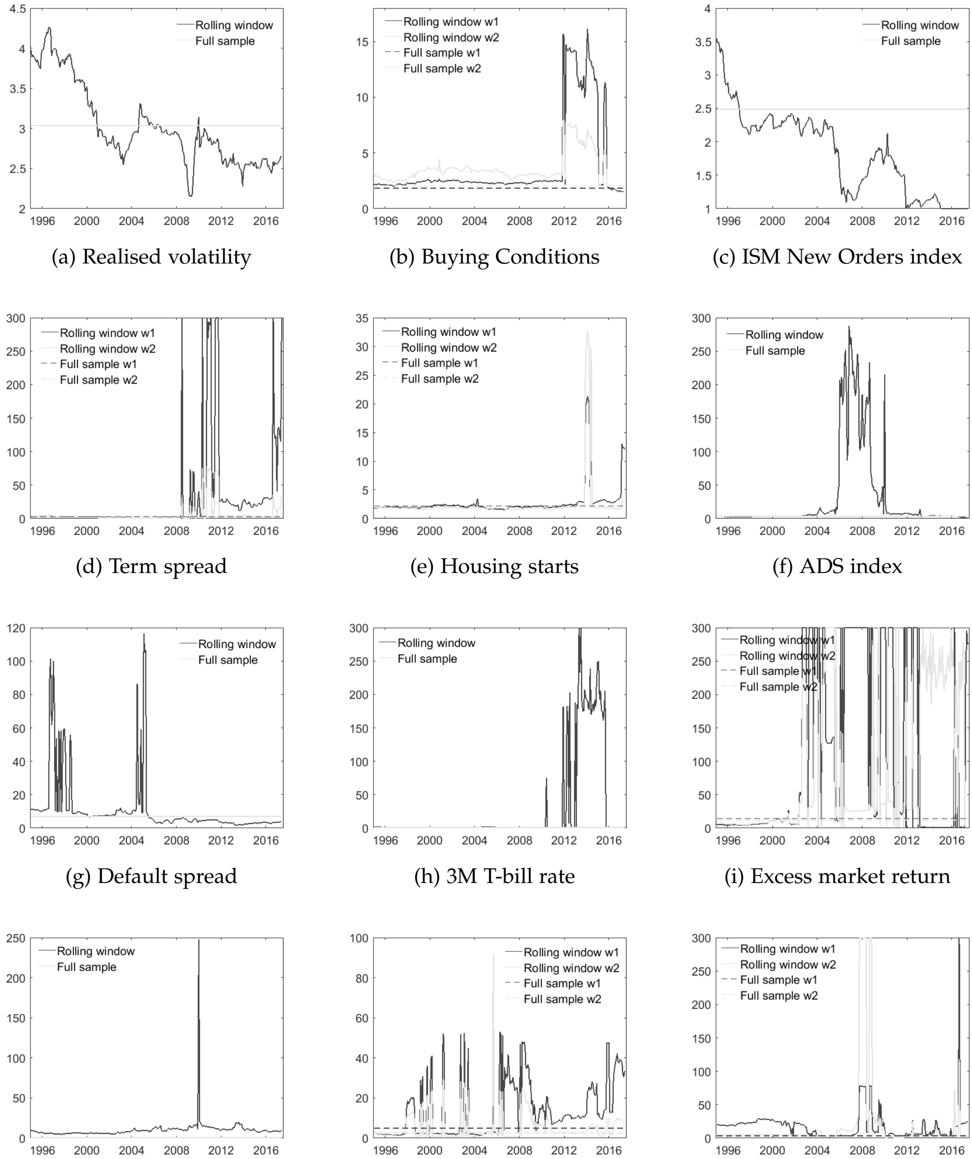
(d) Term spread (e) Housing starts (f) ADS index

(g) Default spread (h) 3M T-bill rate (i) Excess market return

(j) Principal component 1 (k) Principal component 2 (l) Principal component 3

Figure 3.4: Full-sample and rolling window estimates of θ. Dashed lines mark 95% confidence bands.

the results I re-estimate the models with weight parameters (ω_1 and ω_2) fixed at their full-sample values and use an expanding window estimation scheme. Appendix 3.F discusses the results in detail. Overall, fixing the weight parameters has little impact



(a) Realised volatility
(b) Buying Conditions
(c) ISM New Orders index
(d) Term spread
(e) Housing starts
(f) ADS index
(g) Default spread
(h) 3M T-bill rate
(i) Excess market return
(j) Principal component 1
(k) Principal component 2
(l) Principal component 3

Figure 3.5: Full-sample and rolling window estimates of w.

on the forecast accuracy of the models, with only the models driven by excess returns and third PC gaining clearly from fixing the weighting scheme. Thus the instability in the weighting schemes, evident from Figure 3.5, does not tend to adversely impact forecasting performance. The expanding estimation window leads to worse in-sample fit but more stable and more strongly significant parameter estimates. It also leads to more accurate forecasts, especially for financial data. For macroeconomic data and the principal components the evidence is less clear as the mean squared forecast error ratios slightly favour the rolling window estimation scheme. Whether a rolling or expanding window estimation scheme is preferred depends on the explanatory variable, the forecasting horizon as well as the loss function. In most cases the difference is relatively small. The remainder of the paper uses a rolling estimation window as the forecast comparison methods utilised require that a limited memory estimator is used.

3.6 Out-of-sample results

This section begins by discussing the forecasting performance over the whole out-of-sample period, establishing a benchmark. Then, Section 3.6.2 looks at how relative forecasting performance has changed over time, while Section 3.6.3 considers whether forecasting performance varies with the economic environment.

3.6.1 Forecasting performance over the full sample

The results in Table 3.3 indicate that the GJR-GARCH(1,1) model is hard to beat, at least in a statistically significant way.[23] Although macroeconomic data, the term spread and the PCs tend to improve forecasts over longer horizons, the differences are statistically significant only for the term spread and the second PC when absolute forecast errors are used. Other financial variables fail to improve on the benchmark model at any horizon and in fact perform clearly worse in some cases. This is contrary to results using predictive regressions (see, for example, Christiansen et al. (2012)) and could reflect the fact that financial data fail to robustly extract a long-term trend of volatility, which is crucial for the GARCH-MIDAS model. The Mincer-Zarnowitz regressions in Appendix 3.G show that the forecasts are mostly unbiased, although the bias increases as the forecast horizon grows.

3.6.2 Time-variation in relative forecasting performance

We saw in the previous section that many models forecast on average roughly equally well. For example, on the 3M horizon the MSFE ratio for the ISM New Orders index, the term spread and the second and third PCs all equal one. However, this can either be because forecasting performance is similar across models in all time periods or there could be time-variation in relative performance which cancels out over time. To formally investigate time-variation in relative forecasting performance I use the Fluctuation test by Giacomini and Rossi (2010). Figure 3.6 draws the scaled difference in loss functions of a GARCH-MIDAS model and the GJR-GARCH(1,1) model (the test statistic, see equation 3.5) together with two-sided confidence bands.[24] For clarity, I focus

[23] These results echo those in Lindblad (2017).

[24] I set $\alpha = 0.1$ (significance level). I use a Newey-West estimator of the asymptotic variance matrix with lag length $l = 5$, based on the rule-of-thumb, $l = 0.75\ {}^{1/3}\sqrt{T} = 4.77$. The results are robust to changing the lag length to 4 or 8 (results are available upon request). See Giacomini and Rossi (2010) for details on the

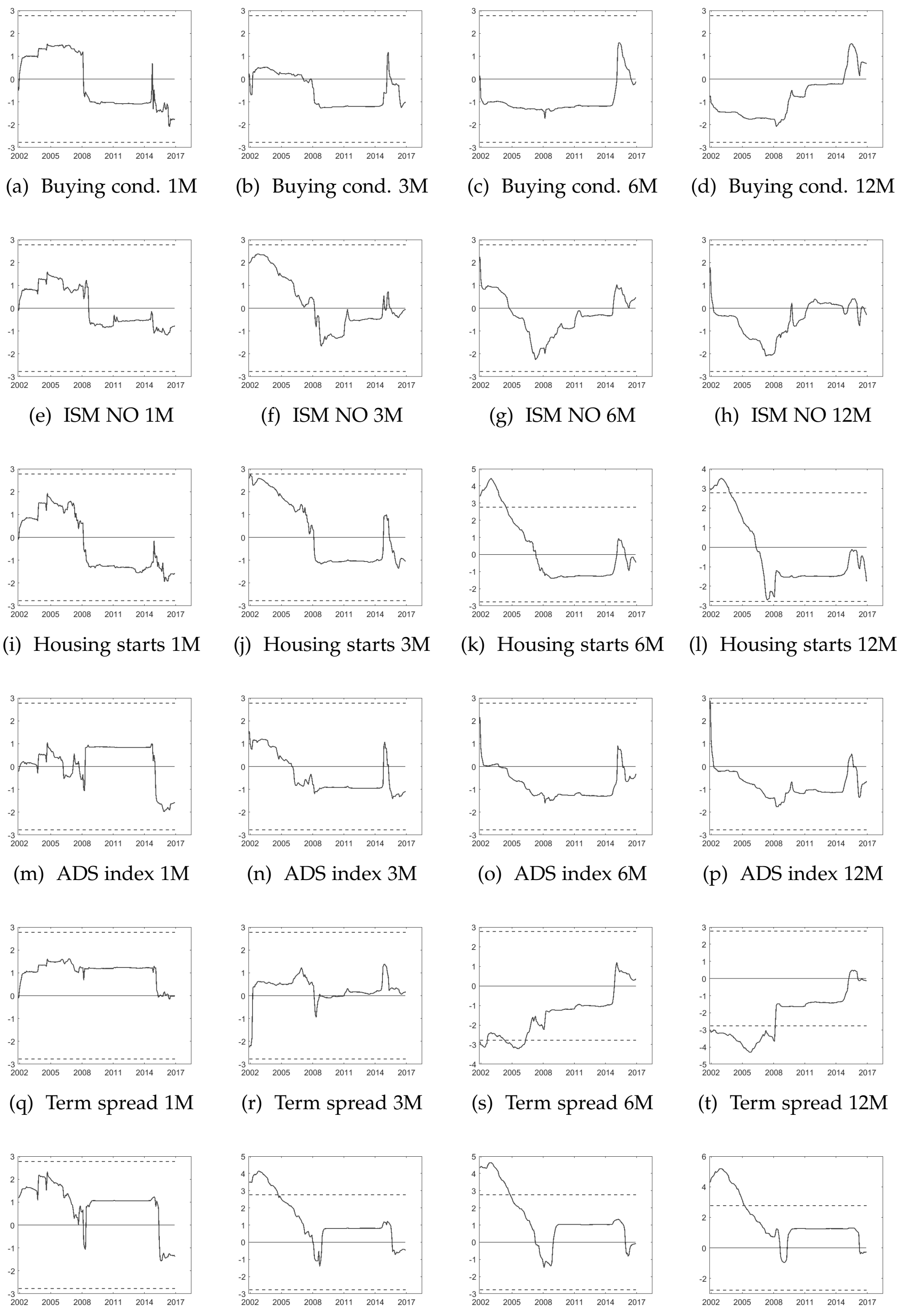
Table 3.3: Full-sample forecast accuracy of the GARCH-MIDAS-X model

	1 month ahead		3 months ahead		6 months ahead		12 months ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Buying Conditions index	1.00	0.98	0.96*	0.91	0.96	0.92	0.98	0.98
ISM New Orders index	0.99	1.00	0.99	1.00	0.99	1.00	0.97	0.99
Housing starts	0.99	0.98	1.00	0.96	0.97	0.93	0.94	0.90
ADS index	1.03	1.12	1.00	0.96	0.99	0.96	1.00	0.98
Term spread	1.03	1.22	0.99	1.00	0.91***	0.94	0.87***	0.91
Default spread	1.09	1.23	1.14	1.12	1.20	1.21	1.20	1.26
3M T-bill rate	1.01	1.04*	1.02	1.04***	1.02	1.04***	1.01	1.03**
Excess market return	1.08	1.36*	1.02	1.01	1.04**	1.02	1.08**	1.04
Realised volatility (RV)	1.14*	1.24	1.22*	1.28	1.29	1.43	1.31	1.49
First PC	0.99	0.88	1.00	0.96	1.02	0.97	1.02	0.98
Second PC	1.01	1.04	0.98	1.00	0.96*	0.98	0.95**	0.97
Third PC	1.05	1.21	0.99	1.00	0.98	0.98	0.99	1.00

Benchmark: GJR-GARCH(1,1). MAFE ratio: $\frac{MAFE_{GMX}}{MAFE_{GARCH}}$*, where* $MAFE_{GMX}$ *stands for the mean absolute forecast error of the GARCH-MIDAS-X model. MSFE ratios calculated equivalently. A value below 1 means the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test.*

on a representative subset of the results with the full results available in Appendix 3.E. Each row corresponds to one economic variable, while the first column presents results for the 1 month forecasting horizon, the second column for the 3M horizon, the third one for the 6M horizon and the rightmost column for the 12M horizon. As the test statistic is calculated over a rolling 6.5 year window the first period is January 1996 - June 2002 and the last period covers January 2011 to June 2017. If the test statistic (solid blue line) exceeds the upper bound (dashed line) the GARCH-MIDAS model produces significantly worse forecasts than the GJR-GARCH(1,1) model, if it drops below the lower bound (dashed line) then the loss of the GJR-GARCH(1,1) model significantly exceeds the loss of the GARCH-MIDAS model. As long as the test statistic is negative the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model, indicating the explanatory variable might be useful for forecasting volatility.

test and the confidence bands.



(a) Buying cond. 1M (b) Buying cond. 3M (c) Buying cond. 6M (d) Buying cond. 12M

(e) ISM NO 1M (f) ISM NO 3M (g) ISM NO 6M (h) ISM NO 12M

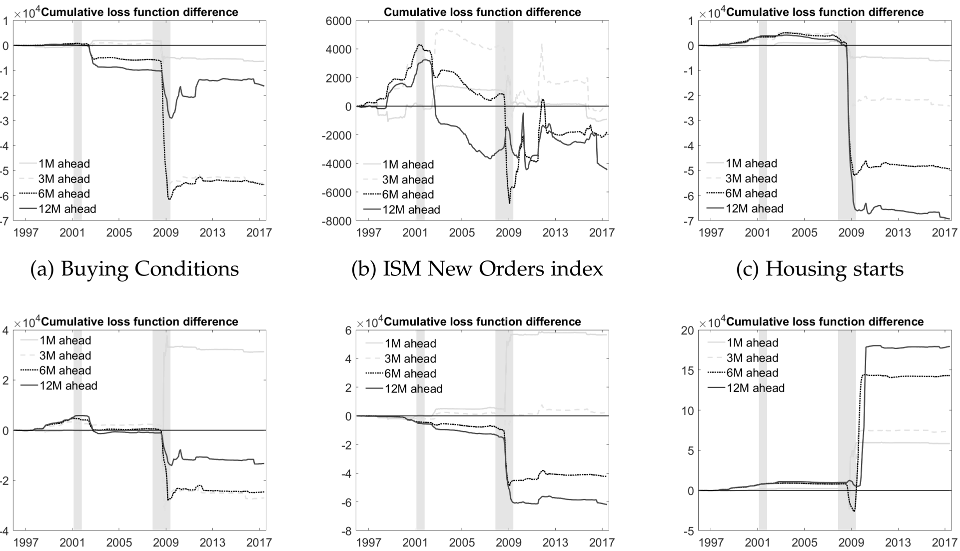
(i) Housing starts 1M (j) Housing starts 3M (k) Housing starts 6M (l) Housing starts 12M

(m) ADS index 1M (n) ADS index 3M (o) ADS index 6M (p) ADS index 12M

(q) Term spread 1M (r) Term spread 3M (s) Term spread 6M (t) Term spread 12M

(u) Default spread 1M (v) Default spread 3M (w) Default spread 6M (x) Default spread 12M

Figure 3.6: Fluctuation test results for loss function differences between the GARCH-MIDAS model and the GJR-GARCH(1,1) model. Squared forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



(a) Buying Conditions (b) ISM New Orders index (c) Housing starts

(d) ADS index (e) Term spread (f) Default spread

Figure 3.7: Cumulative sum of loss function differences between the GARCH-MIDAS models and the GJR-GARCH(1,1) model ($Loss^2_{MIDAS} - Loss^2_{GARCH}$). An upward sloping segment thus indicates the GJR-GARCH model outperforms the GARCH-MIDAS model. Grey areas mark NBER dated US recessions.

Overall, the forecast accuracy of the different GARCH-MIDAS models vary significantly over time, with the differences in performance becoming larger as the forecasting horizon increases. There is, however, no model that is superior on all forecasting horizons. In general, the GJR-GARCH model significantly outperforms some of the GARCH-MIDAS models early in the sample period, such as those driven by housing starts and the default spread. On the other hand, the term spread driven GARCH-MIDAS model outperforms the benchmark in a statistically significant way over the 6 and 12 month horizons in the first half of the out-of-sample period. In addition, for several of the macroeconomic variables the test statistic is primarily negative over the longer forecasting horizons, implying macroeconomic data can be useful for forecasting. The benefit of augmenting a basic GARCH model with financial data (excluding the term spread) remains weak. Looking at the end of the sample, the performance of economic data has improved on the one month horizon and housing starts is the best predictor on the 12 month horizon.[25]

To see in detail how relative forecasting performance has evolved over time Figure 3.7 draws the cumulative sum of loss function differences for the four forecasting horizons. The relative performance of many GARCH-MIDAS models clearly improves during the recession and financial crisis in 2007-2008, especially for horizons longer

[25]When using mean absolute forecasts errors the results are stronger in terms of statistical significance in favour of the macroeconomic data, and the recent relative performance of the GARCH-MIDAS models is more convincing, see Appendix 3.D.

than one month. On the other hand, for most of the financial data performance significantly weakens during the same time period. For the ISM New Orders index forecasting performance over long horizons improves especially between and during the two recessions. Overall we can conclude that when forecasting over short horizons (1M ahead) the standard GARCH model is quicker to adapt to recessions or financial crises and therefore tends to produce more accurate forecasts, whereas the long-term trend component is useful when forecasting over longer horizons.

To conclude, whereas Conrad and Loch (2014) found that macroeconomic variables (which partly overlap with my data set) significantly improved forecasts (compared to the GARCH-MIDAS-RV model) between the two recessions and since mid-2008, strong evidence in favour of using macroeconomic data cannot be found here, despite many macroeconomic variables performing well especially during the recession around 2008-2009. As discussed earlier, the diverging conclusions arise from, for example, differences in evaluation window length and benchmark model.

3.6.3 Effect of the economic environment on forecast accuracy

As shown above, the ability of economic data to predict long-term stock return volatility varies over time. However, is this purely random variation, or can it be explained by the economic or market environment? As discussed in, for example, Hamilton and Lin (1996), it is logical to assume that the dynamic behaviour of the economy is different during expansions and contractions, and that the business cycle can thus be broken down into two distinct states. When forecasting volatility it is also plausible that the volatility environment can affect relative forecast accuracy.

I divide the out-of-sample period into sub-samples according to a business cycle (or volatility) indicator and compare forecasting performance separately for recession (or high volatility) and expansion (or low volatility) periods.[26] If there are differences in forecasting performance between, for example, recession and expansion periods, then the knowledge of whether we are, or are approaching, a recession or expansion period could help us choose a more accurate forecasting model.

3.6.3.1 Business cycles

I first divide the sample into positive and negative growth periods based on the sign of industrial production growth. As a robustness check I also divide the data based on the NBER dated US recessions (see Appendix 3.H).

From Table 3.4 we can see that the GJR-GARCH(1,1) model is still difficult to outperform at short horizons. Macroeconomic variables, the term spread and the principal components do, however, improve forecasts in negative growth periods and in particular over long horizons. This is in line with the results in Figure 3.7, where many of the macroeconomic variables improved forecast in particular during the latest recession. The GARCH-MIDAS model augmented by the term spread is also the best model in expansions over the longest horizon. The performance of GARCH-MIDAS models driven by other financial data remains weak, regardless of the business cycle. The main conclusions carry over to the MAFEs (Table 3.11) and to using NBER recession

[26]See Appendix 3.H for a classification of the regimes.

Table 3.4: Effect of business cycle (IP growth) on forecasting performance

	1 month ahead		3 months ahead		6 months ahead		12 months ahead	
	> 0	< 0	> 0	< 0	> 0	< 0	> 0	< 0
Buying Conditions index	1.00	0.97	1.01	0.89	1.03	0.90	1.07	0.96*
ISM New Orders index	1.00	0.99	1.03	0.99	1.05	0.99	0.99	0.99
Housing starts	1.01	0.96	1.03	0.95	1.05	0.91	1.01	0.88
ADS index	0.99	1.18	1.01	0.94	1.05	0.95	1.04	0.97
Term spread	1.03	1.30	1.01	1.00	1.00	0.93	0.96	0.90
Default spread	1.02	1.31	1.50	1.05	2.44	1.00	2.52	1.05
3M T-bill rate	1.04*	1.05	1.08**	1.03**	1.10**	1.03**	1.11**	1.01
Excess market return	0.98	1.51*	1.03	1.01	0.99	1.02	1.13	1.02
Realised volatility (RV)	1.02	1.33	1.56	1.22	2.85	1.18	3.69	1.11
First PC	1.00	0.83	1.00	0.96	1.13	0.95	1.08	0.97
Second PC	1.04	1.04	1.07	0.98	1.09	0.96	1.04	0.98
Third PC	1.04	1.29	1.04	0.99	1.10	0.98	1.11	0.99

*Benchmark: GJR-GARCH(1,1) model. MSFE ratio: $\frac{MSFE_{GMX}}{MSFE_{GARCH}}$, where $MSFE_{GMX}$ stands for the mean squred forecast error of the GARCH-MIDAS-X model. A value below 1 means the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test. Positive (> 0) and negative (< 0) growth months defined according to the sign of annualised monthly industrial production growth (manufacturing only, most recent value): 95 low growth and 163 high growth periods. $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$.*

periods instead (Table 3.16), and these results are even stronger in terms of statistical significance.

3.6.3.2 Volatility environment

I next divide the sample based on the VIX index, and as a robustness check (see Appendix 3.H) the St. Louis Fed Financial Stress Index (STLFSI)[27], in order to determine how forecast accuracy of GARCH-MIDAS models is impacted by the volatility environment.

Many of the economic variables are useful for forecasting volatility in low volatility periods while the gains are less clear in high volatility periods (Table 3.5). The improvements in forecasting performance in low volatility periods are in many cases statistically significant, even on the one month horizon, for which also financial data is useful. The low volatility periods take place right before the financial crisis in 2007-2008 and after roughly 2013 (see Figure 3.37). Thus the results in this section confirm that the differences in forecasting performance uncovered in Section 3.6.2 can at least partly be explained by changes in the volatility environment. Interestingly, the model driven by the 3M T-bill rate clearly improves forecasts in low volatility periods while leading to significantly worse forecasts in the high volatility periods. Especially the second and third principal components perform very well in low volatility environments. It seems intuitive that economic data is more important for forecasts during

[27]The STLFSI consists of 18 series, including interest rates, yield curves and the VIX index.

Table 3.5: Effect of volatility environment on forecasting performance

	1 month ahead		3 months ahead		6 months ahead		12 months ahead	
	Low	High	Low	High	Low	High	Low	High
Buying Conditions index	0.98	0.98	0.72	0.91	0.71**	0.92	0.73	0.98
ISM New Orders index	0.92**	1.00	0.70	1.01	0.72**	1.00	0.64**	1.00
Housing starts	1.01	0.98	0.75	0.96	0.86	0.93	0.74	0.90
ADS index	0.98	1.13	0.83	0.96	0.87	0.97	0.83	0.98
Term spread	1.00	1.23	0.72	1.01	0.58***	0.94	0.37***	0.92
Default spread	0.98	1.23	0.90	1.12	1.03	1.22	3.66	1.21
3M T-bill rate	0.95**	1.05*	0.71*	1.04***	0.72***	1.04***	0.61***	1.04***
Excess market return	0.90*	1.37*	0.86	1.02	1.26*	1.01	1.56	1.03
Realised volatility (RV)	0.78**	1.26	0.72	1.29	0.97	1.44	3.80	1.44
First PC	0.92*	0.88	0.81	0.97	0.93	0.97	1.00	0.98
Second PC	0.91*	1.04	0.59**	1.00	0.51***	0.99	0.39***	1.00
Third PC	0.97	1.22	0.69*	1.01	0.68**	1.00	0.53***	1.01

*Benchmark: GJR-GARCH(1,1) model. MSFE ratio: $\frac{MSFE_{GMX}}{MSFE_{GARCH}}$, where $MSFE_{GMX}$ stands for the mean squared forecast error of the GARCH-MIDAS-X model. A value below 1 means the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test. High / low volatility months are classified according to the median of the VIX index: 147 months of high volatility and 111 months of low volatility. $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$.*

calm markets, while the GARCH model, which reacts more quickly to changes in the market environment, performs better in high volatility environments. The main results are robust to using mean absolute forecast errors (Table 3.12).[28] Clearly, if we could correctly anticipate being in a low volatility environment we might be able to improve volatility forecasts by including economic data in a GARCH model.

3.7 Conditional predictive ability

In the previous section I determined that relative forecasting performance depends on the business cycle and the volatility environment. This section builds on this insight and explores whether relative forecasting performance is predictable using information on the state of the economy and the volatility environment available at the forecast origin.[29] This information could be exploited in forecast combination schemes or forecast model selection. I apply the conditional predictive ability test by Giacomini and White (2006) to statistically test whether relative forecasting performance is predictable using the (expected) state of the business cycle (real-time professional recession probabili-

[28]The results do not get strong support from MSFE ratios when dividing the sample based on the financial stress index (Table 3.17), indicating low volatility rather than low stress is important.

[29]An alternative approach would be to let the effect of the economic data (i.e., parameter θ) depend on the economic environment directly. Appendix 3.I explores this alternative. The results indicate that the relationship between volatility and the economic variables strengthens when the economic environment is weak or financial conditions are tight, but overall the differences in forecasting performance are small.

ties by the Survey of Professional Forecasters[30]) and an indicator for financial market volatility (VIX index). The interpretation of the test is such that if we find that the conditional test rejects (Table 3.6) while the unconditional test fails to reject (Table 3.3), then even though average performance is roughly equal, relative performance could have been predicted using information on the economic and market environment at the forecast origin.[31] On the other hand, if the unconditional test rejects while the conditional test does not, then the conditional test could have low power or the unconditional test could be undersized (see Giacomini and White (2006) for details).

Table 3.6: Conditional predictive ability test

	1 month ahead		3 months ahead		6 months ahead		12 months ahead	
	Abs	Sq	Abs	Sq	Abs	Sq	Abs	Sq
Buying Conditions index	0.52	0.69	0.96	0.63	0.88*	0.67	0.60**	0.81
ISM New Orders index	0.88	0.59	0.69	0.34	0.77	0.50	0.71	0.68
Housing starts	0.52	0.59	0.64	0.65	0.70**	0.64	0.92*	0.93
ADS index	0.55	0.52	0.62	0.64	0.67	0.63	0.55	1.00
Term spread	0.52	0.47	0.73	0.54	1.00***	0.68	1.00***	0.94
Default spread	0.56	0.48	0.48	0.55	0.41	0.48	0.44	0.43
3M T-bill rate	0.32	0.35	0.29	0.00	0.43	0.01**	0.43*	0.00**
Excess market return	0.58	0.48	0.56	0.41	0.01	0.46	0.13*	0.01
Realised volatility	0.58	0.57	0.55	0.56	0.50	0.53	0.47	0.49
First PC	0.68	0.67	0.67	0.69	0.56	0.65	0.50	0.68*
Second PC	0.44	0.45	1.00	0.50	0.90	0.52	0.93*	0.98
Third PC	0.45	0.45	0.74	0.33	0.76	0.58	0.57	0.54

*The entries are the proportions of times the predictive loss difference ($\hat{\delta}_n' h_t$), as outlined in Giacomini and White (2006), indicates that the GARCH-MIDAS-X model is better than the GJR-GARCH(1,1) model. A number over 0.50 indicates the GARCH-MIDAS model is chosen over 50% of the time, and vice versa. *, ** and *** indicate a rejection of the null hypothesis of equal conditional predictive ability (over the full sample) at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test. Conditioning variables: level of the VIX index and recession probabilities from the Survey of Professional Forecasters. Test function: $h_t = [1\ v_t]$, where v_t is the conditioning information. For the SPF data I use: 1Q ahead recession probability for 1M and 3M ahead forecasts, 2Q ahead for 6M, and 4Q ahead for 12M ahead forecasts. Abs refers to using the absolute value as the loss function, while Sq refers to using squared forecast errors.*

The results for the conditional predictive ability test over the full sample indicate that the forecast errors are predictable over long horizons for, for example, models including the Buying Conditions index, housing starts or the term spread, but the results depend on the loss function used (Table 3.6). However, comparing the significance of the loss function differences to those in Table 3.3 it is evident that conditioning on the recession probabilities and the VIX index we can only find some modest improvements in predictability for the models including the Buying Conditions index or housing starts.

[30] Quarterly data is transformed into monthly frequency by keeping it fixed within each quarter.

[31] The tests were also performed with each of the conditioning variables separately, but this did not bring significant new insights compared to the results here. The results are largely robust to using the STLFSI, NBER recession dates and industrial production growth. The results are available upon request.

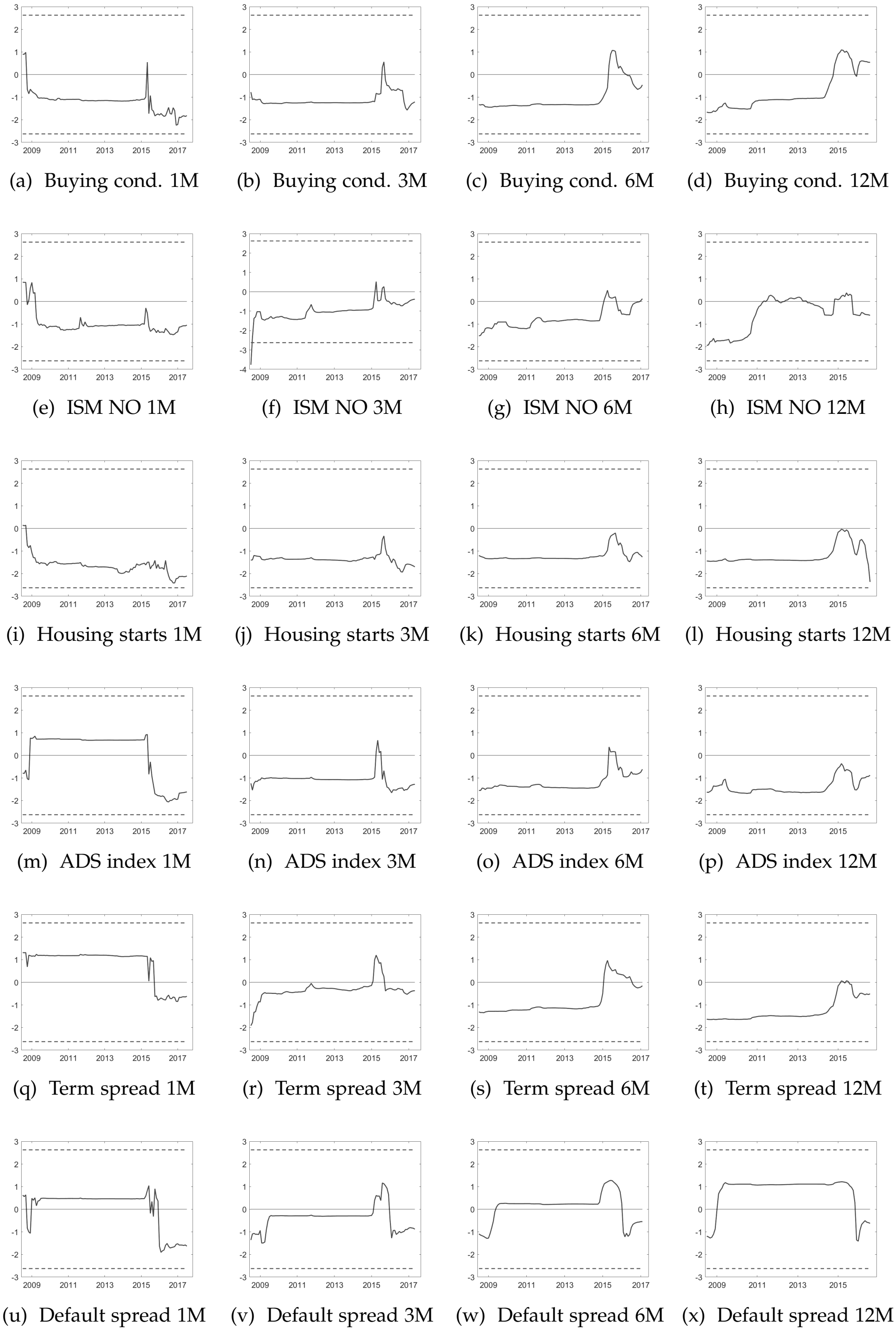
I also determine whether using a simple decision rule based on conditional predictive ability, following the general idea outlined in Section 4 of Giacomini and White (2006), could help improve forecasting performance. In a first step, in each out-of-sample period, I regress the loss function difference on the test function (i.e., the conditioning information, h_t) over the out-of-sample period and denote the resulting regression coefficients by $\hat{\delta}'_n$. In a second step, if equal conditional predictive ability is rejected, I choose the forecast from the GARCH-MIDAS model if the out-of-sample predictive loss difference is negative ($\hat{\delta}'_n h_t < 0$) and the GJR-GARCH(1,1) model forecast if $\hat{\delta}'_n h_t > 0$. If the conditional predictive ability test cannot distinguish between the two forecasts, I take an equal-weighted average of the two model forecasts. The forecast series produced are therefore hybrids of the two individual model forecasts and their forecast combination. I implement the decision rule over an expanding window with an initial size of six years, conditioning on the SPF forecasts and the VIX index.[32]

Table 3.7: Decision rule based on conditional predictive ability

	1M ahead		3M ahead		6M ahead		12M ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Buying Conditions index	**0.99**	0.98	0.96**	0.95	0.95**	0.94	**0.97**	**0.97**
ISM New Orders index	**0.98****	1.00	**0.96*****	0.99	**0.95*****	0.99	**0.95*****	0.99*
Housing starts	0.99	0.99*	**0.97****	0.96	**0.93****	0.91	0.91	0.89
ADS index	**1.01**	**1.04**	**0.98**	0.97	0.98	0.97	**0.98**	0.98*
Term spread	**1.00**	**1.11**	**0.96***	**0.99**	0.91***	**0.93**	0.86***	0.91
Default spread	**1.02**	**1.03**	**0.95**	**0.98**	**0.97**	**1.02**	**0.99**	**1.06**
3M T-bill rate	**1.00**	**1.02**	**0.98****	**1.00**	**0.97****	**1.00**	**0.97***	**1.00**
Excess market return	**1.03**	**1.16**	**1.02**	**1.01**	**1.01**	**1.01**	**1.02***	**1.02**
Realised volatility (RV)	**1.06**	**1.04**	**1.07**	**1.04**	**1.16**	**1.10**	**1.19**	**1.13**
First PC	**0.97**	**0.84**	**0.95*****	0.96*	**0.95****	**0.95**	**0.97****	**0.96**
Second PC	**1.00**	**1.02**	0.97***	**0.98**	0.94***	**0.97**	0.92***	0.98
Third PC	**1.02**	**1.09**	**0.97****	0.99	**0.96****	0.99	**0.95****	**0.99**

*Benchmark: GJR-GARCH(1,1) model. MAFE ratio: $\frac{MAFE_{DR}}{MAFE_{GARCH}}$, where $MAFE_{DR}$ stands for the mean absolute forecast error of the decision rule based forecast series. MSFE ratios calculated equivalently. A value below 1 means the decision rule based forecast outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal unconditional predictive ability (over the full sample) at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test. Conditioning variable and test function: see notes on Table 3.6. Calculated over an expanding window, with an initial size of six years. Bold numbers indicate the forecast error of the decision rule based forecast is smaller than the forecast error of the underlying GARCH-MIDAS model.*

[32]The first window is therefore January 1996 - December 2001.



(a) Buying cond. 1M (b) Buying cond. 3M (c) Buying cond. 6M (d) Buying cond. 12M

(e) ISM NO 1M (f) ISM NO 3M (g) ISM NO 6M (h) ISM NO 12M

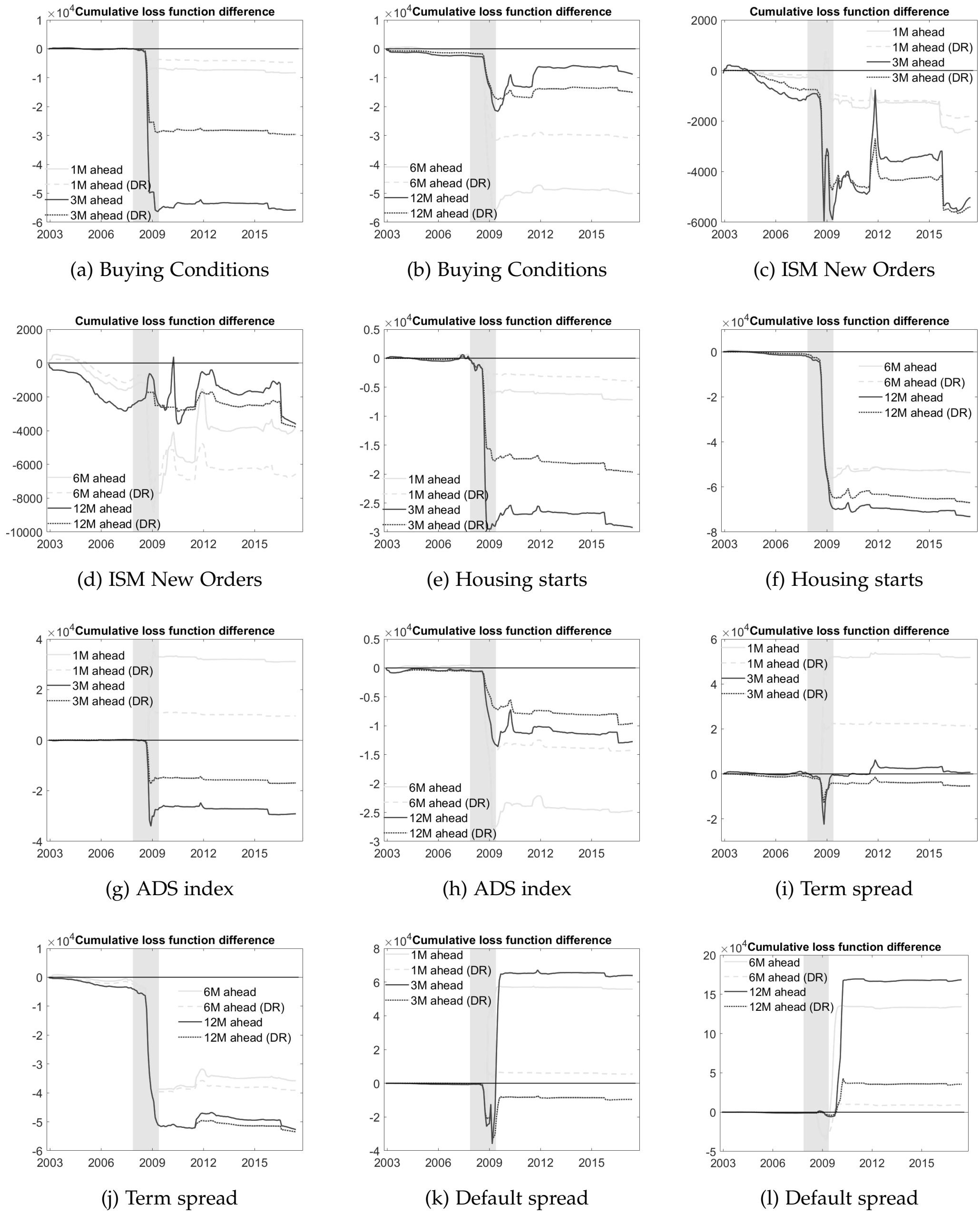
(i) Housing starts 1M (j) Housing starts 3M (k) Housing starts 6M (l) Housing starts 12M

(m) ADS index 1M (n) ADS index 3M (o) ADS index 6M (p) ADS index 12M

(q) Term spread 1M (r) Term spread 3M (s) Term spread 6M (t) Term spread 12M

(u) Default spread 1M (v) Default spread 3M (w) Default spread 6M (x) Default spread 12M

Figure 3.8: Fluctuation test results for loss function differences between the decision rule based forecasts and the GJR-GARCH(1,1) model. Squared forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



(a) Buying Conditions (b) Buying Conditions (c) ISM New Orders

(d) ISM New Orders (e) Housing starts (f) Housing starts

(g) ADS index (h) ADS index (i) Term spread

(j) Term spread (k) Default spread (l) Default spread

Figure 3.9: Cumulative sum of loss function differences between the decision rule (DR) based forecasts and the GJR-GARCH(1,1) model ($Loss^2_{DR} - Loss^2_{GARCH}$). An upward sloping segment thus indicates the GJR-GARCH model outperforms the decision rule. Grey areas mark NBER dated US recessions.

Table 3.6 reports the share of times the out-of-sample predictive loss difference ($\hat{\delta}_n' h_t$) of the decision rule outlined above indicates the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model.[33] A share over 0.50 indicates the GARCH-MIDAS model is chosen more often. The shares vary clearly between variables and horizons, but tend to be over 0.50 for models incorporating macroeconomic explanatory variables and less than 0.50 for financial data, in line with earlier results.

Turning to the hybrid forecasts based on the decision rule outlined above, we can see that they lead to modest forecast improvements (Table 3.7) on forecasting horizons longer than 1 month. Many of the decision rule based forecasts produce more accurate forecasts than the benchmark GJR-GARCH model, and several of the decision rules based on macroeconomic variables or principal components now produce forecasts which are significantly better than the benchmark, in particular when using the absolute value loss function. However, only in a subset of cases does the decision rule also improve on the performance of the underlying GARCH-MIDAS model (bold numbers which are also less than one in the table). The clearest improvements seem to be for the ISM New Order index, housing starts and the principal components based hybrid forecasts.

Based on the Fluctuation test (Figures 3.8 and 3.26) the decision rule results in forecasts which consistently outperform the GJR-GARCH model, but the differences are not statistically significant.[34] Looking at the cumulative sums of the loss function differences (Figure 3.9) we can see that the differences in forecasting performance between the individual GARCH-MIDAS models and the decision rule based hybrid forecasts occur mainly during, or immediately after, the latest recession. For the macroeconomic data the decision rule mainly dampens their forecast gains, while for the financial data the decision rules can lead to substantial forecast improvements during this time. Clearly, a simple decision rule like the one outlined above does bring some advantages compared to individual models, but it still fails to completely utilise the differences in time-varying forecasting performance.

The cumulative loss function differences (Figure 3.9) show that the largest gains from using the decision rule occur for financial data, for which the deterioration in forecasting performance occurring especially in conjunction with the recession and financial crisis is moderated. On the other hand, the averaging of forecasts in the decision rule also tends to dampen the good performance of the macroeconomic data around the recession. A better way of predicting relative forecast accuracy could clearly further enhance performance.

3.8 Forecast combination schemes

If the relative forecasting performance of models varies over time, forecast combination methods can be useful. The seminal paper by Bates and Granger (1969) already

[33]Contrary to the forecast combinations based on the decision rule in Table 3.7, I do not require statistical significance of the conditional predictive ability test when calculating the shares. These numbers are intended as a general guideline only for how often the predictive loss differences imply the GARCH-MIDAS model outperforms the baseline GARCH model.

[34]Note that due to the initial calculations required for the decision rule the forecast comparisons begin six years later than those in Section 3.6.

concluded that combination forecasts can outperform the individual forecasts, a conclusion widely confirmed in later literature.[35] In practice, simple forecast combination methods, such as equal weights, often lead to more accurate forecasts than more complicated schemes (for example, Clemen (1989)).

This section explores combining the GARCH-MIDAS model forecasts using both simple and time-varying combination schemes. Because the financial variables produced clearly inferior forecasts on all horizons over most time periods, I focus on combining the forecasts produced by the macroeconomic variables and the term spread. The simple combination schemes are the mean, the median and the trimmed mean[36] of the GARCH-MIDAS forecasts. The time-varying alternatives either use time-varying weights (the discounted mean square or absolute prediction error (DMSPE or DMAPE) following Stock and Watson (2004)) or choose the forecast(s) to be included by ranking the forecasts based on past performance (i.e., past forecast errors), similar to, for example, Aiolfi and Timmermann (2006). The DMSPE forecast combination scheme is used by, for example, Rapach et al. (2010) for equity premium prediction and Paye (2012) for stock market volatility forecasts in the predictive regression setting.

The combination forecasts are weighted averages of the N individual volatility forecasts ($\hat{\sigma}^2_{i,t+1}$): $\hat{\sigma}^2_{c,t+1} = \Sigma_{i=1}^{N}\omega_{i,t}\hat{\sigma}^2_{i,t+1}$, where the weights depend on the chosen combination method. For example, the simple mean combination puts $\omega_{i,t} = \frac{1}{N}$. The DMAPE weights depend on the historical performance of the models:

$$\omega_{i,t} = \frac{\phi_{i,t}^{-1}}{\Sigma_{j=1}^{N}\phi_{j,t}^{-1}}, \text{ where } \phi_{i,t} = \Sigma_{s=1}^{t-h}\eta^{t-h-s}|\sigma^2_{s+h} - \hat{\sigma}^2_{i,s+h}|$$

and h is the forecasting horizon.[37] $0 < \eta \leq 1$ is the discount factor: $\eta = 1$ is the basic case from Bates and Granger (1969) for uncorrelated individual forecasts. When $\eta < 1$ recent forecast accuracy is weighted more heavily. I use $\eta = 1$ and $\eta = 0.5$. Stock and Watson (2004) conclude that for macroeconomic forecasting more discounting ($\eta = 0.9$) usually performs at least no better than less discounting (for example, $\eta = 1$).

If there is clear persistence in forecasting performance and the differences between model accuracy are large, we can potentially improve on the simple mean by excluding the worst performing models in each period. Considering the results in the previous section it is clear that there were some models which produced inferior forecasts for a prolonged period of time, and preselecting the included forecasting models based on past performance can thus be beneficial. I rank the forecasting models in each period and for each horizon based on average past performance over an expanding window. In each out-of-sample period I then pick the forecast of the model that has had the best average forecasting performance up until the forecast origin ('Previously best'), as well as take the mean of the forecasts of the best-performing three models ('Mean (best three)').

Table 3.8 gives the mean absolute and squared forecast error ratios of the combination forecasts compared to the benchmark GJR-GARCH model. Over the 1 month horizon performance of the combination forecasts is similar, or slightly worse, to the bench-

[35] See, for example, Clemen (1989), Chan et al. (1999) and Stock and Watson (1999).

[36] The trimmed mean refers to removing the smallest and the largest forecasts each period and taking a mean of the remaining forecasts.

[37] The DMSPE weights are calculated equivalently, except that squared forecast errors are used.

Table 3.8: Full-sample forecast accuracy of forecast combinations

	1 month ahead		3 months ahead		6 months ahead		12 months ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Mean	1.00	1.04	0.96**	0.96	0.95**	0.94	0.94**	0.95
Median	0.99	0.99	0.97*	0.96	0.96*	0.94	0.94***	0.96*
Trimmed mean	0.99	1.01	0.97*	0.96	0.95*	0.94	0.94***	0.95
DMAPE / DMSPE, $\eta = 1$	1.00	1.04	0.96**	0.96	0.95**	0.94	0.94***	0.95
DMAPE / DMSPE, $\eta = 0.5$	1.00	1.04	0.96**	0.95	0.94**	0.94	0.93**	0.94
Previously best	1.05	1.13	0.98	0.91	0.91***	0.94	0.87***	0.91*
Mean (best three)	1.00	1.02	0.97	0.94	0.94***	0.94	0.93***	0.95

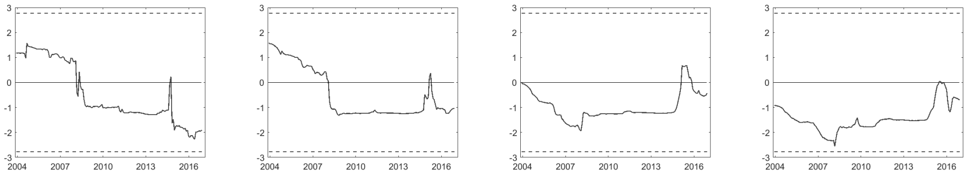
*Benchmark: GJR-GARCH(1,1) model. MAFE ratio: $\frac{MAFE_{combo}}{MAFE_{GARCH}}$, where $MAFE_{combo}$ stands for the mean absolute forecast error of the combination forecast using the method stated in the first column. MSFE ratios calculated equivalently. A value below 1 means the combination forecast outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test. DMAPE is used for the MAFE ratios and the DMSPE for the MSFE ratios. The last four combination schemes are based on forecasting performance over an expanding window with an initial size of 12 months. Due to initial calculations all forecast comparisons are for the period January 1998 - June 2017 (234 periods).*

mark. Over longer horizons the forecast combinations tend to outperform the benchmark contrary to most individual forecasts. The differences are statistically significant especially for absolute forecast errors. The Fluctuation test, which tests whether the forecasting performance is time-varying, reveals that the test statistics are, especially on horizons longer than 3 months, predominantly negative, and the GJR-GARCH(1,1) never significantly outperforms any of the combination forecasts (see Figure 3.10).[38] Thus the combination forecasts are superior to most of the individual forecasts by more consistently outperforming the benchmark model, although the differences tend not to be statistically significant. The exception is the combination scheme using the forecast of the best performing model, which on longer horizons largely replicates the performance of the term spread driven GARCH-MIDAS model. Overall the difference between simple and time-varying combination methods is small.[39] When considering the absolute value of the forecast errors (Figure 3.22) we can see that not only have the combination forecasts outperformed the benchmark in a statistically significant way over long horizons in the early part of the sample, but in many cases the forecast combinations also perform clearly better towards the end of the forecasting sample. To shed further light on the performance of the combination forecasts I draw the cumulative sum of the loss function differences in Figure 3.11. They highlight the usefulness of combining GARCH-MIDAS forecasts during the latest recession for forecasting horizons above one month.

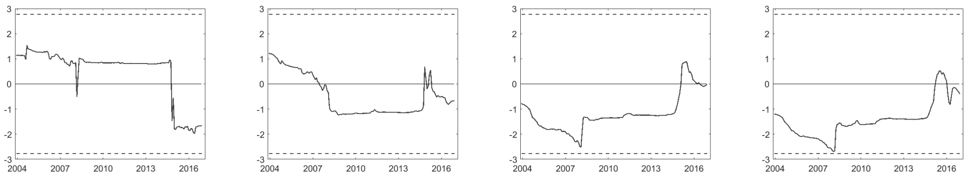
Comparing the forecast combinations to the principal component driven models (see Appendix 3.E) reveals that forecast combinations tend to perform better than the models using information from a large amount of economic data. Therefore, it seems

[38]As the forecast combination schemes produce very similar forecasts, I only include a representative subset of the results in the main text. Full results are available in Appendix 3.E.

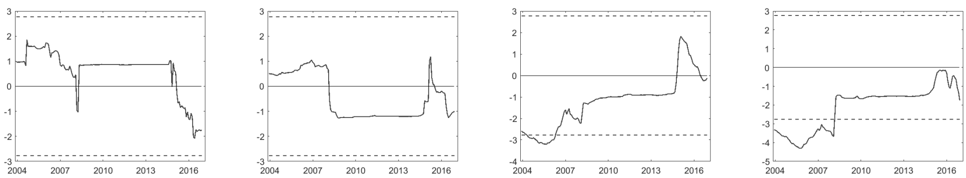
[39]The number of models being combined is modest (5), and a larger amount of individual models could reveal larger differences between the different combination schemes.



(a) Median 1M (b) Median 3M (c) Median 6M (d) Median 12M

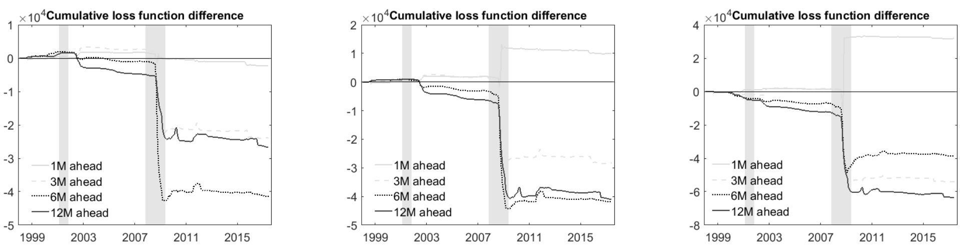


(e) DMSPE 1M (f) DMSPE 3M (g) DMSPE 6M (h) DMSPE 12M



(i) Prev. best 1M (j) Prev. best 3M (k) Prev. best 6M (l) Prev. best 12M

Figure 3.10: Fluctuation test results for loss function differences between the forecast combinations of the GARCH-MIDAS models and the GJR-GARCH(1,1) model. Squared forecast errors. Dashed lines represent 90% confidence bands. DMSPE with $\eta = 0.5$. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



(a) Median (b) DMSPE ($\eta = 0.5$) (c) Previously best

Figure 3.11: Cumulative sum of loss function differences between the forecast combinations of GARCH-MIDAS models and the GJR-GARCH(1,1) model ($Loss^2_{combo} - Loss^2_{GARCH}$). An upward sloping segment thus indicates the GJR-GARCH model outperforms the GARCH-MIDAS model. Grey areas mark NBER dated US recessions.

that for forecasting volatility using GARCH-MIDAS models pooling models is more useful than pooling information. Overall, forecast combinations seem useful for forecasting long-term volatility in many periods and provide forecasts that are consistently at least slightly better than the benchmark model.

3.9 Conclusion

This paper evaluates the time-variation in relative forecasting performance of models for stock return volatility, with focus on using macroeconomic and financial data to enhance long-horizon volatility forecasts. The paper addresses three related questions. First, how stable are the rolling window parameter estimates of the GARCH-MIDAS models and how does this instability affect model fit and forecasting performance? Second, how does the predictive ability of different economic variables vary over time, and is this time-variation linked to the economic or financial market environment? Third, can forecast accuracy be improved by utilising the time-variation in forecasting performance or by combining individual GARCH-MIDAS forecasts?

A rolling window estimation scheme leads to more volatile parameter estimates but also better in-sample fit, when compared to an expanding window estimation scheme. Whether a rolling or expanding window is preferred for out-of-sample forecasting depends on the explanatory data, forecasting horizon and loss function. When forecasting over long horizons there are clear shifts in relative forecasting performance over time implying that (time-varying) forecast combination methods could be useful. Macroeconomic variables improve predictions especially in low volatility periods but also in periods of weak economic growth, while financial data driven GARCH-MIDAS models struggle to outperform the benchmark GJR-GARCH model, with the exception of short horizon forecasts in low volatility environments. The largest gains over the past two decades from using macroeconomic data – and forecast combinations – were realised during the latest recession and financial crisis. Only some forecast errors are predictable conditioning on, for example, the volatility environment, and therefore only modest gains are realised using a decision rule based on conditional forecasting performance.

It is clear that no single forecasting model or combination scheme excels on all horizons and in all time periods, although the decision rules produce forecasts which modestly, but consistently, outperform the benchmark. When forecasting 12 months ahead the term spread driven GARCH-MIDAS model and the forecast combinations perform well, while the decision rule and some of the GARCH-MIDAS-X models do well over the medium term. Over the 1 month horizon there is some evidence that macroeconomic data is useful for forecasting. However, with the exception of the term spread, these differences in forecasting performance tend not to be statistically significant, especially when using squared forecast errors. As the GJR-GARCH model is rarely significantly better than the GARCH-MIDAS models and never significantly outperforms the combination forecasts or the decision rule forecasts, it does, however, seem beneficial to use economic data for long-horizon volatility forecasting.

References

Aiolfi, M. and A. Timmermann (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics 135*, 31–53.

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*(4), 885–905.

Andreou, E., E. Ghysels, and A. Kourtellos (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics 158*, 246–261.

Asgharian, H., C. Christiansen, and A. J. Hou (2015). Effects of macroeconomic uncertainty on the stock and bond markets. *Finance Research Letters 13*, 10–16.

Asgharian, H., A. J. Hou, and F. Javed (2013). The importance of the macroeconomic variables in forecasting stock return variance: a GARCH-MIDAS approach. *Journal of Forecasting 32*, 600–612.

Awartani, B. M. A. and V. Corradi (2005). Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries. *International Journal of Forecasting 21*, 167–183.

Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance 61*(4), 1645–1680.

Bates, J. M. and C. W. Granger (1969). The combination of forecasts. *Or*, 451–468.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Cakmakli, C. and D. van Dijk (2010). Getting the most out of macroeconomic information for predicting stock returns and volatility. *Tinbergen Institute Discussion paper, No. 10-115/4*.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal 101*, 157–179.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies 1*, 195–228.

Chan, Y. L., J. H. Stock, and M. W. Watson (1999). A dynamic factor model framework for forecast combination. *Spanish Economic Review 1*, 91–121.

Chauvet, M. and S. Potter (2013). Forecasting output. In *Handbook of Economic Forecasting*, Volume 2, pp. 141–194. Elsevier.

Christiansen, C., M. Schmeling, and A. Schrimpf (2012). A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics 27*(6), 956–977.

Christoffersen, P. F. and F. X. Diebold (2000). How relevant is volatility forecasting for financial risk management? *The Review of Economics and Statistics 82*(1), 12–22.

Clark, T. and K. West (2006). Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics 135*, 155–186.

Clark, T. E. and M. W. McCracken (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics 105*(1), 85–110.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting 5*, 559–581.

Conrad, C. and K. Loch (2014). Anticipating long-term stock market volatility. *Journal of Applied Econometrics 30*(7), 1090–1114.

Davis, N. and A. M. Kutan (2003). Inflation and output as predictors of stock returns and volatility: international evidence. *Applied Financial Economics 13*(9), 693–700.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*, 253–263.

Diebold, F. X. and K. Yilmaz (2008). Macroeconomic volatility and stock market volatility, world-wide. *NBER Working Paper* (14269).

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica 50*(4), 987–1008.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *The Review of Economics and Statistics 95*(3), 776–797.

Engle, R. F. and G. Lee (1999). *A permanent and transitory component model of stock return volatility, R. Engle and H. White (ed.) Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*. Oxford, UK: Oxford University Press.

Engle, R. F. and J. G. Rangel (2008). The SPLINE-GARCH model for low frequency volatility and its global macroeconomic causes. *Review of Financial Studies 21*, 1187–1222.

Errunza, V. and K. Hogan (1998). Macroeconomic determinants of European stock market volatility. *European Financial Management 4*(3), 361–377.

Ghysels, E., A. Rubia, and R. Valkanov (2009). Multi-period forecasts of volatility: Direct, iterated and mixed-data approaches. *working paper*.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The MIDAS touch: Mixed data sampling regression models. *EFA 2009 Bergen Meetings Paper*.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2005). There is a risk-return tradeoff after all. *Journal of Financial Economics* (76), 509–548.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics 131*(1-2), 59–95.

Ghysels, E., A. Sinko, and R. Valkanov (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Giacomini, R. and B. Rossi (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics 25*, 595–620.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*, 1545–1578.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Hamilton, J. D. and G. Lin (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics 11*(5), 573–593.

Hansen, P. R. and A. Lunde (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics 20*, 873–889.

Lindblad, A. (2017). Sentiment indicators and macroeconomic data as drivers for long-term stock market volatility. *HECER Discussion Paper* (413).

McCracken, M. W. (2000). Robust out-of-sample inference. *Journal of Econometrics 99*(2), 195–223.

McCracken, M. W. and S. Ng (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics 34*(4), 574–589.

Mincer, J. A. and V. Zarnowitz (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, pp. 3–46. NBER.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics 160*, 246–256.

Paye, B. S. (2012). Deja vol: Predictive regressions for aggregate stock market volatility using macroeocnomic variables. *Journal of Financial Economics 106*, 527–546.

Pierdzioch, C., J. Dopke, and D. Hartmann (2008). Forecasting stock market volatility with macroeconomic variables in real time. *Journal of Economics and Business 60*, 256–276.

Poon, S.-H. and C. W. Granger (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature 41*, 478–539.

Rapach, D. E., J. K. Strauss, and G. Zhou (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies 23*(2), 821–862.

Schwert, W. G. (1989). Business cycles, financial crises, and stock volatility. *Carnegie-Rochester Conference Series on Public Policy 31*, 83–126.

Stock, J. H. and M. W. Watson (1999). Forecasting inflation. *Journal of Monetary Economics 44*, 293–375.

Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: the role of asset prices. *Journal of Economic Literature 41*, 788–829.

Stock, J. H. and M. W. Watson (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting 23*(6), 405–430.

Wang, F. and E. Ghysels (2015). Econometric analysis of volatility component models. *Econometric Theory 31*, 362–393.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica 64*, 1067–1084.

Appendices

3.A Data description

- CRSP index, excess market return: Kenneth French's Data Library
- ISM New Orders index: FRED database and the Institute for Supply Management (https:https://www.instituteforsupplymanagement.org/)
- Buying Conditions index: the University of Michigan consumer confidence report (https:https://data.sca.isr.umich.edu/)
- Housing starts, industrial production growth: Philadelphia Fed real time center
- ADS index: Philadelphia Fed real time center. For details see https://www.philadelphiafed.org/research-and-data/real-time-center
- Survey of Professional Forecaster data, real-time professional recession probabilities: Philadelphia Fed real time center
- Interest rates, default spread, VIX index, Chicago Fed Adjusted National Financial Conditions (ANFCI) index, St. Louis Fed Financial Stress Index (STLFSI): FRED database (St. Louis Fed)
- NBER recession dates: NBER (http://www.nber.org/cycles.html)
- FRED-MD Database:
  https://research.stlouisfed.org/econ/mccracken/fred-databases/

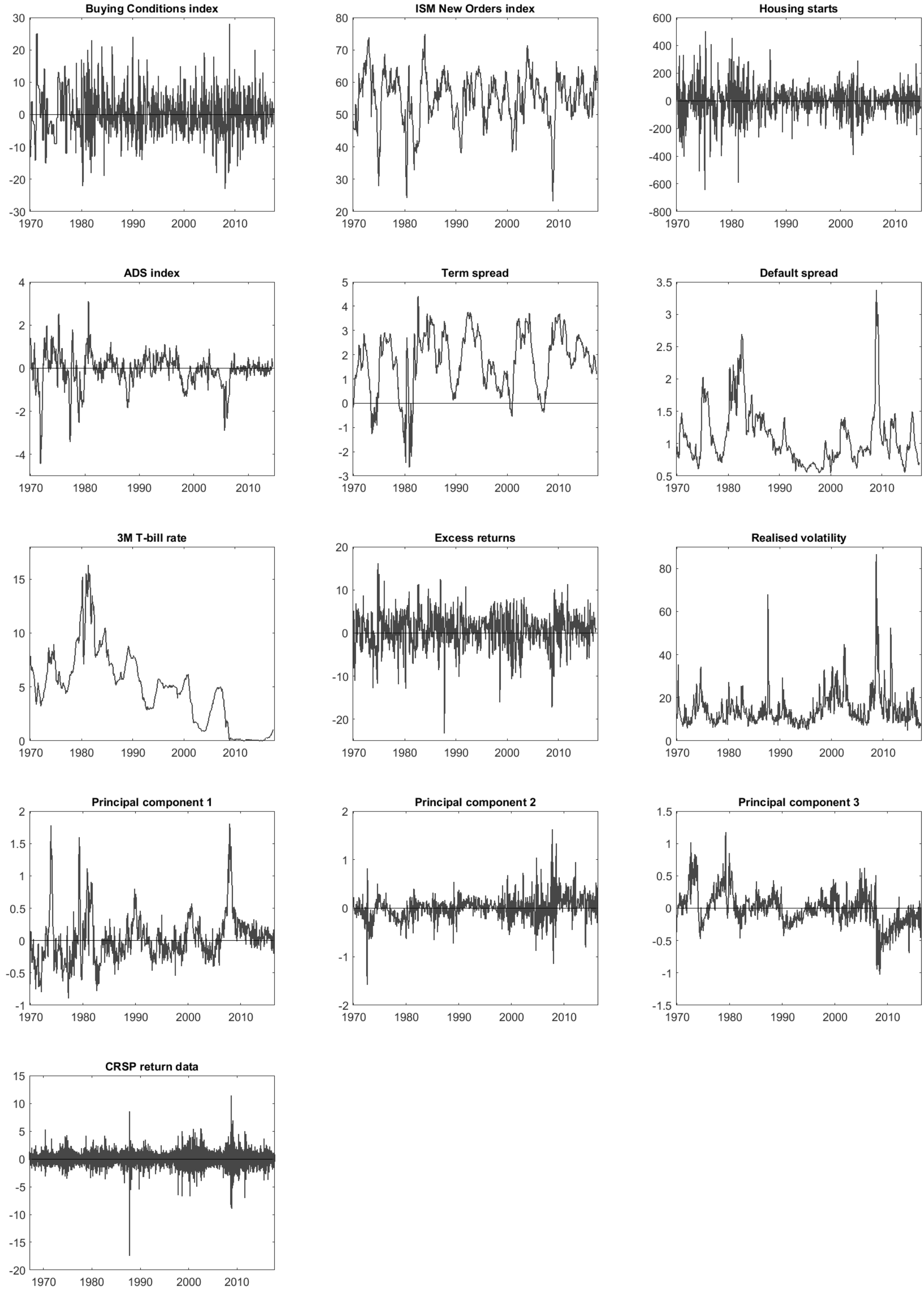


Figure 3.12: Explanatory data and return data.

3.B Principal Component Analysis

This appendix presents the ten highest marginal R^2s for the first four factors, extracted from the FRED-MD Database (full-sample). See McCracken and Ng (2016) for details on the data and the methodology to calculate the PCs. The numbers in parentheses denote the marginal R^2 for each factor, i.e., how much each factor explains of the overall variation in the data.

Table 3.9: Ten highest marginal R^2s for the first four factors

PC 1 (0.1472)		PC 3 (0.0685)	
Employment: Goods-Prod. Industries (USGOOD)	0.7106	Moody's Aaa Corporate Bond - Fed Funds (AAAFFM)	0.4487
Total nonfarm employment (PAYEMS)	0.7101	10Y Treasury C - Fed Funds (T10YFFM)	0.4438
IP: Manufacturing (SIC) (IPMANSICS)	0.6888	Moody's Baa Corporate Bond - Fed Funds (BAAFFM)	0.4333
IP Index (INDPRO)	0.6552	5Y Treasury C - Fed Funds (T5YFFM)	0.3956
Employment: Manufacturing (MANEMP)	0.6512	3M Treasury C - Fed Funds (TB3SMFFM)	0.3290
IP: Final Products and Nonindustrial Supplies (IPFPNSS)	0.6116	6M Treasury C - Fed Funds (TB6SMFFM)	0.3118
Employment: Durable goods (DMANEMP)	0.6001	1Y Treasury C - Fed Funds (T1YFFM)	0.2648
Capacity Utilization (manufacturing) (CUMFNS)	0.5927	CPI: Commodities (CUSR0000SAC)	0.2467
IP: Final Products (Market Group) (IPFINAL)	0.5137	Pers. Cons. Exp: Nondur. goods (DNDGRG3M086SBEA)	0.2437
IP: Durable Materials (SRVPRD)	0.4803	CPI (excl. shelter) (CUUR0000SA0L2)	0.2383
PC 2 (0.0708)		**PC 4 (0.0558)**	
CPI: Commodities (CUSR0000SAC)	0.5680	1Y Treasury Rate (GS1)	0.5073
Personal Cons. Exp. (Nondur.) (DNDGR3M086SBEA)	0.5573	5Y Treasury Rate (GS5)	0.4922
CPI (excl. shelter) (CUUR0000SA0L2)	0.5441	Moody's Seasoned Aaa Corporate Bond Yield (AAA)	0.4830
CPI: All Items (CPIAUCSL)	0.5321	6M Treasury Bill (TB6MS)	0.4707
CPI (excl. medical care) (CUSR0000SA0L5)	0.5016	10Y Treasury Rate (GS10)	0.4537
Personal Cons. Expenditure: Chain index (PCEPI)	0.4762	Moody's Seasoned Baa Corporate Bond Yield (BAA)	0.4374
CPI: Transportation (CPITRNSL)	0.4702	3M Treasury Bill: (TB3MS)	0.3749
CPI (excl. food) (CPIULFSL)	0.4299	3M AA Financial Commercial Paper Rate (CP3Mx)	0.3749
PPI: Finished Consumer Goods (PPIFCG)	0.3121	New Orders for Consumer Goods (ACOGNO)	0.2009
PPI: Finished goods (PPIFGS)	0.3595	S&P's Comp. Common Stock: Div. Yield (S&P div yield)	0.1864

Sample period: December 1959 - May 2017. Based on the FRED-MD Database, vintage June 2017, by McCracken and Ng (2016).

Figure 3.13 presents rolling window results for the principal components analysis, detailing which series are most often chosen into the first three principal components. The first time-varying PC relates mainly to real activity and employment related series. The second PC mainly relates to interest rates and interest rate spreads but also to price variables. For the third PC one cluster relates to price variables, a second to interest rates and a third one relates to housing market data. See McCracken and Ng (2016) for details on the series and to link specific series numbers to the name of the series.

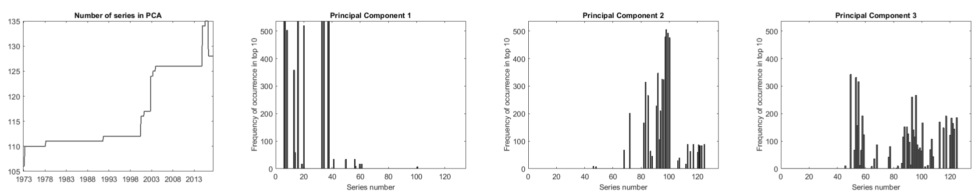


Figure 3.13: Time-varying composition of the first three PCs. First panel shows how the number of series in the data set varies over time.

3.C Restricted versus unrestricted weighting scheme

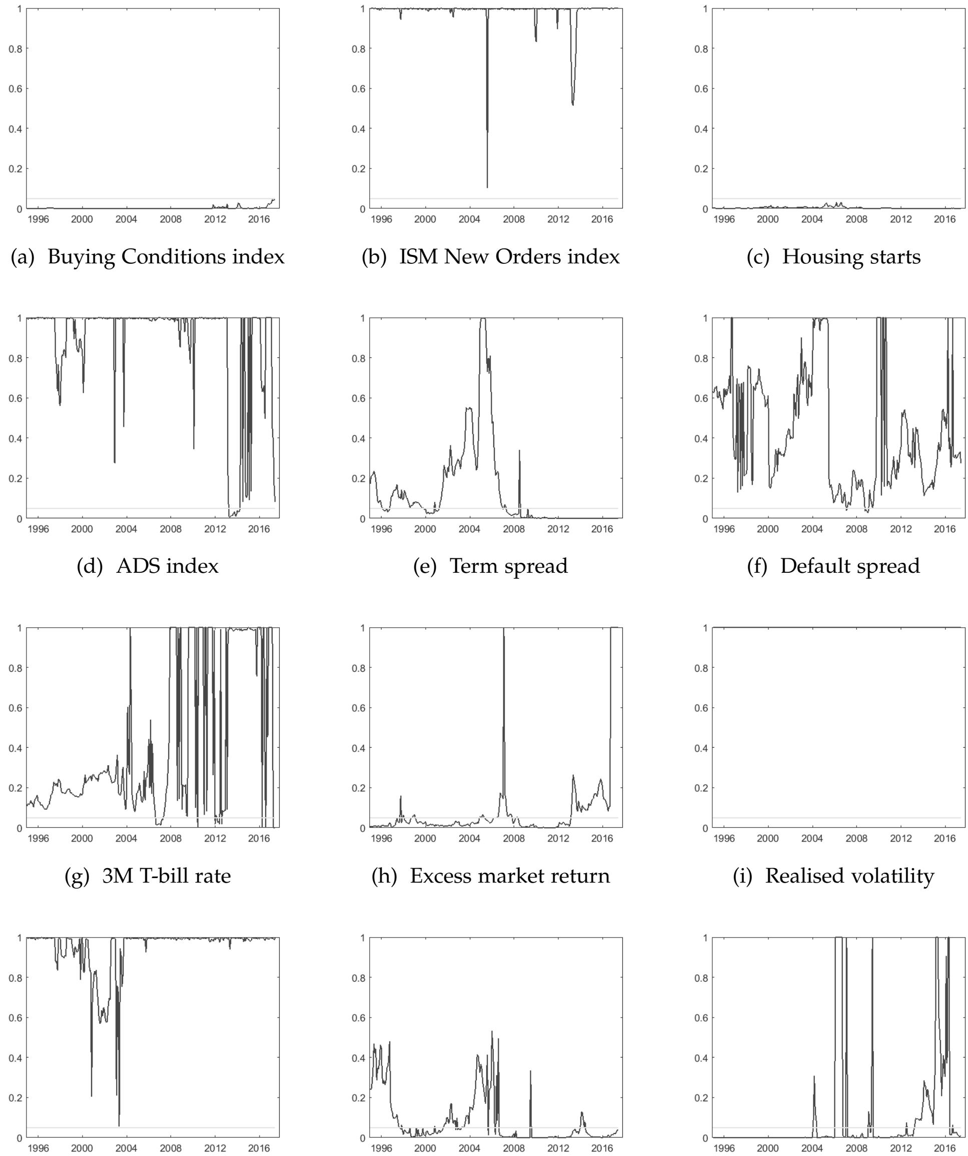
This appendix explores the implications of estimating one or two weights in the MIDAS polynomial. This is especially crucial for the term spread, excess market returns, PC2 and PC3, for which the choice of the optimal weighting schemes varies over time. Figure 3.14 shows the time variation in p-values from a likelihood ratio test between a model with one or two weights for each GARCH-MIDAS model. For the model driven by the term spread or the second PC two weights have been preferred lately, while the opposite is true for the excess market return and the third PC. Figure 3.15 graphs the time-varying estimates for θ, which are mostly similar for the unrestricted and restricted weighting schemes, except when the model with two weights is clearly superior.

Table 3.10: Out-of-sample results: restricted vs. unrestricted models

	1 month ahead		3 months ahead		6 month ahead		12 month ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Term spread	1.00	0.97	1.01	1.02	1.03	1.03	1.03	1.01
3M T-bill rate	1.00	0.99	1.00	1.00	1.00	1.01	1.00	1.00
Excess market return	0.98	1.00	1.01	1.00	0.99	1.00	1.00	1.00
Realised volatility (RV)	0.96	0.84	1.00	0.92	1.02	0.95	1.03	0.98
Second principal component	1.00	0.96	1.01	1.01	1.00	1.00	0.99	0.98
Third principal component	0.97	0.90	0.99	1.01	0.99	1.01	0.97	0.99

MAFE ratio: $\frac{MAFE_{GMX1w}}{MAFE_{GMX2w}}$, where $MAFE_{GMX1w}$ stands for the mean absolute forecast error of the GARCH-MIDAS-X model estimated using a restricted weighting scheme. The MSFE ratios are calculated equivalently. A value below 1 means the GARCH-MIDAS model with a restricted weighting scheme outperforms the GARCH-MIDAS model with an unrestricted weighting scheme. $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$.

Figure 3.16 and Table 3.10 consider how the forecast errors change when the models are estimated with a restricted or an unrestricted weighting scheme. Only the models for which the choice is ambiguous are considered. Over the full sample the differences in forecasting performance tend to be small. Over time the differences in forecasting performance vary. On the 1M horizon the restricted model tends to perform better especially during the recession for most variables. On the other hand, for the term spread and realised volatility the unrestricted model performs better over the 12M horizon. The choice of an unrestricted weighting scheme seems well founded for, for example, the term spread, while for the third principal component the restricted model seems to produce more accurate forecasts on all horizons in many time periods. Overall, the choice between a restricted and unrestricted weighting scheme depends on the forecasting horizon and time period.



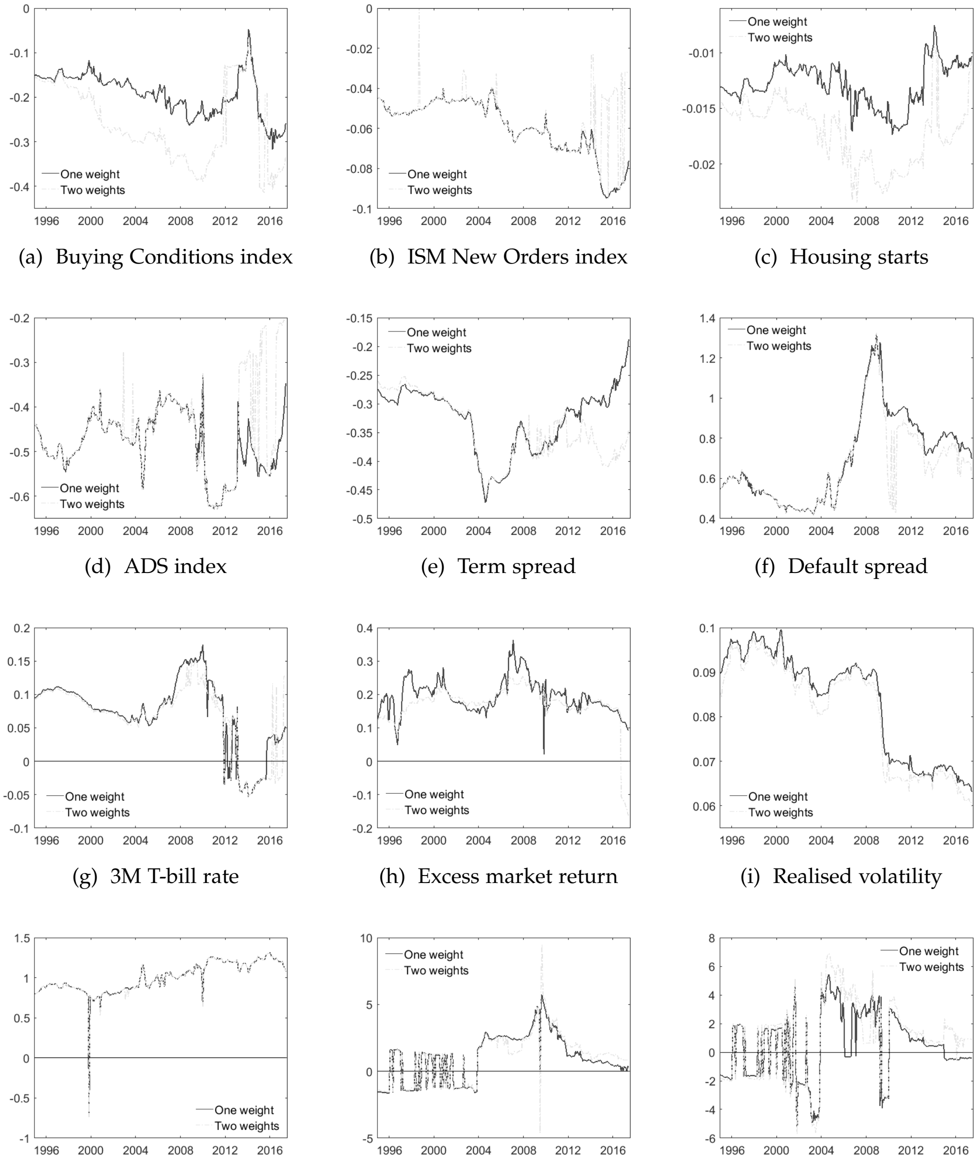
(a) Buying Conditions index (b) ISM New Orders index (c) Housing starts

(d) ADS index (e) Term spread (f) Default spread

(g) 3M T-bill rate (h) Excess market return (i) Realised volatility

(j) Principal component 1 (k) Principal component 2 (l) Principal component 3

Figure 3.14: p-values from a likelihood ratio test between restricted and unrestricted models. Horizontal line indicates 5% significance level.



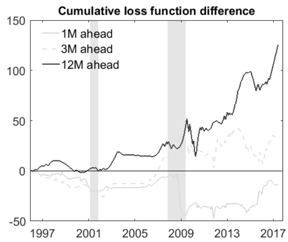
(a) Buying Conditions index (b) ISM New Orders index (c) Housing starts

(d) ADS index (e) Term spread (f) Default spread

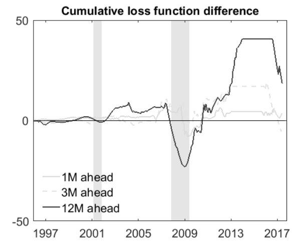
(g) 3M T-bill rate (h) Excess market return (i) Realised volatility

(j) Principal component 1 (k) Principal component 2 (l) Principal component 3

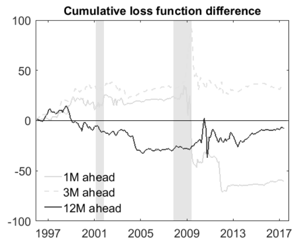
Figure 3.15: Rolling window estimates of θ from the GARCH-MIDAS models estimated using a restricted and an unrestricted weighting scheme.



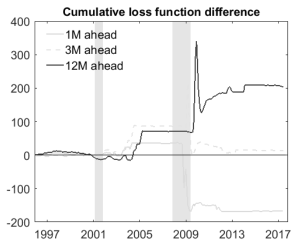
(a) Term spread



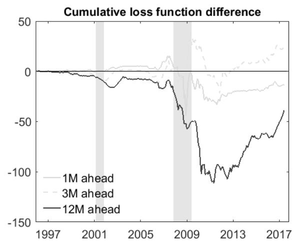
(b) 3M T-bill rate



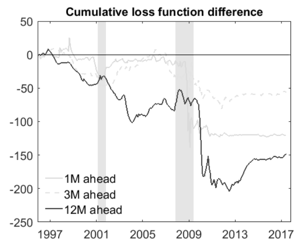
(c) Excess market returns



(d) Realised volatility



(e) Principal component 2

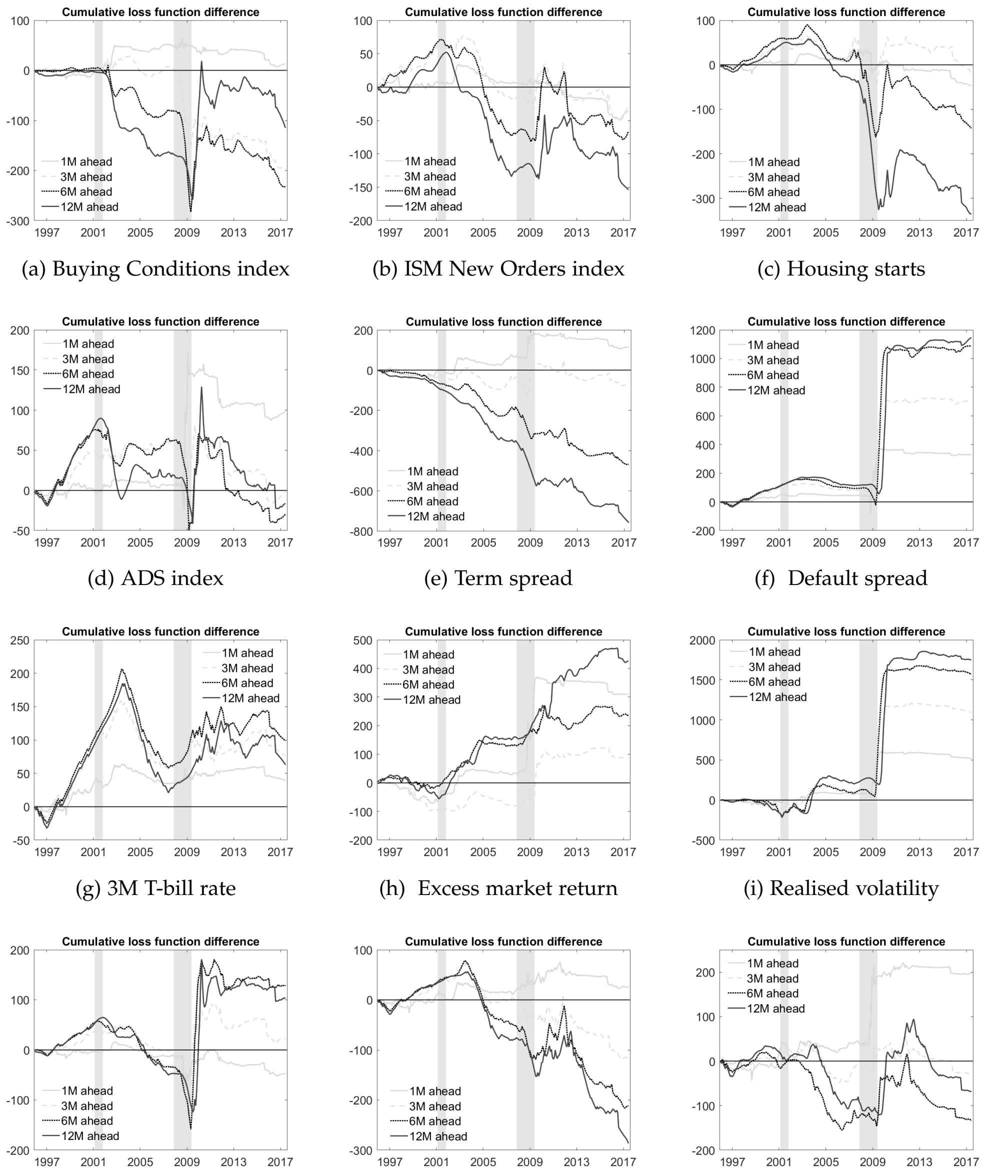


(f) Principal component 3

Figure 3.16: Cumulative sum of loss function differences between models estimated with a restricted and an unrestricted weighting scheme. An upward sloping line indicates the unrestricted model is superior.

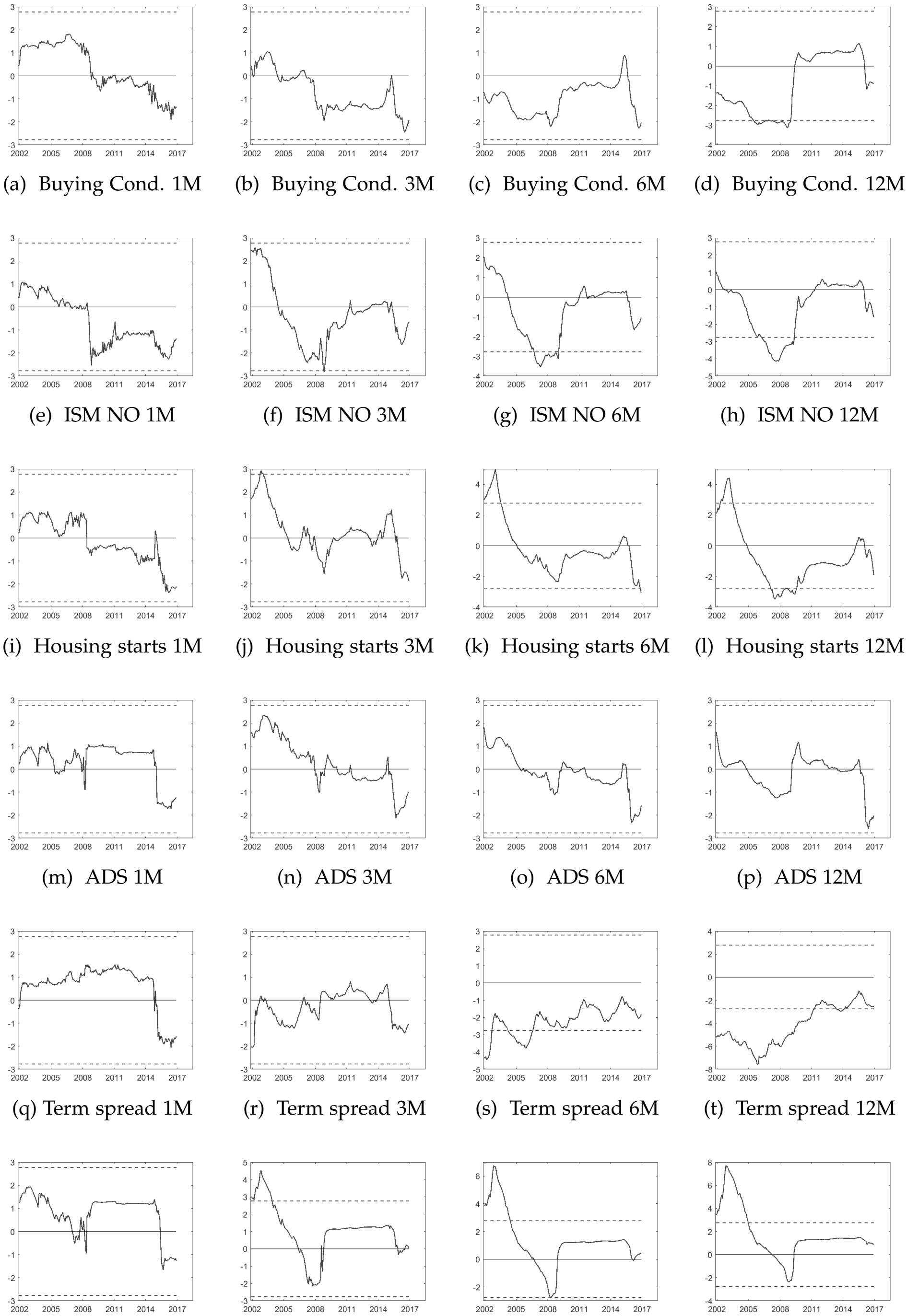
3.D Mean absolute forecast errors

This appendix presents results for the absolute value of the forecast errors, as a robustness check to the squared forecast errors presented in the main text. The absolute forecast error puts less emphasis on individual large forecast errors and therefore has a tendency to display stronger statistical significance, which can be seen clearly in the Fluctuation test by Giacomini and Rossi (2010) (Figures 3.18 and 3.19). The recent performance of the GARCH-MIDAS models has been more convincing when looking at MAFEs than MSFEs. The cumulative forecast errors in Figure 3.17 indicate that many of the macroeconomic variables have been particularly useful between and during the two recessions, while less useful right after the recession period. The MAFEs relating to the remaining MSFE results in the paper are available upon request.



(a) Buying Conditions index
(b) ISM New Orders index
(c) Housing starts
(d) ADS index
(e) Term spread
(f) Default spread
(g) 3M T-bill rate
(h) Excess market return
(i) Realised volatility
(j) Principal component 1
(k) Principal component 2
(l) Principal component 3

Figure 3.17: Cumulative sum of loss function differences between the GARCH-MIDAS models and the GJR-GARCH model ($|Loss_{GMX}| - |Loss_{GARCH}|$). An upward sloping segment thus indicates the GJR-GARCH model outperforms the GARCH-MIDAS model. Grey areas mark NBER dated US recessions.



(a) Buying Cond. 1M (b) Buying Cond. 3M (c) Buying Cond. 6M (d) Buying Cond. 12M

(e) ISM NO 1M (f) ISM NO 3M (g) ISM NO 6M (h) ISM NO 12M

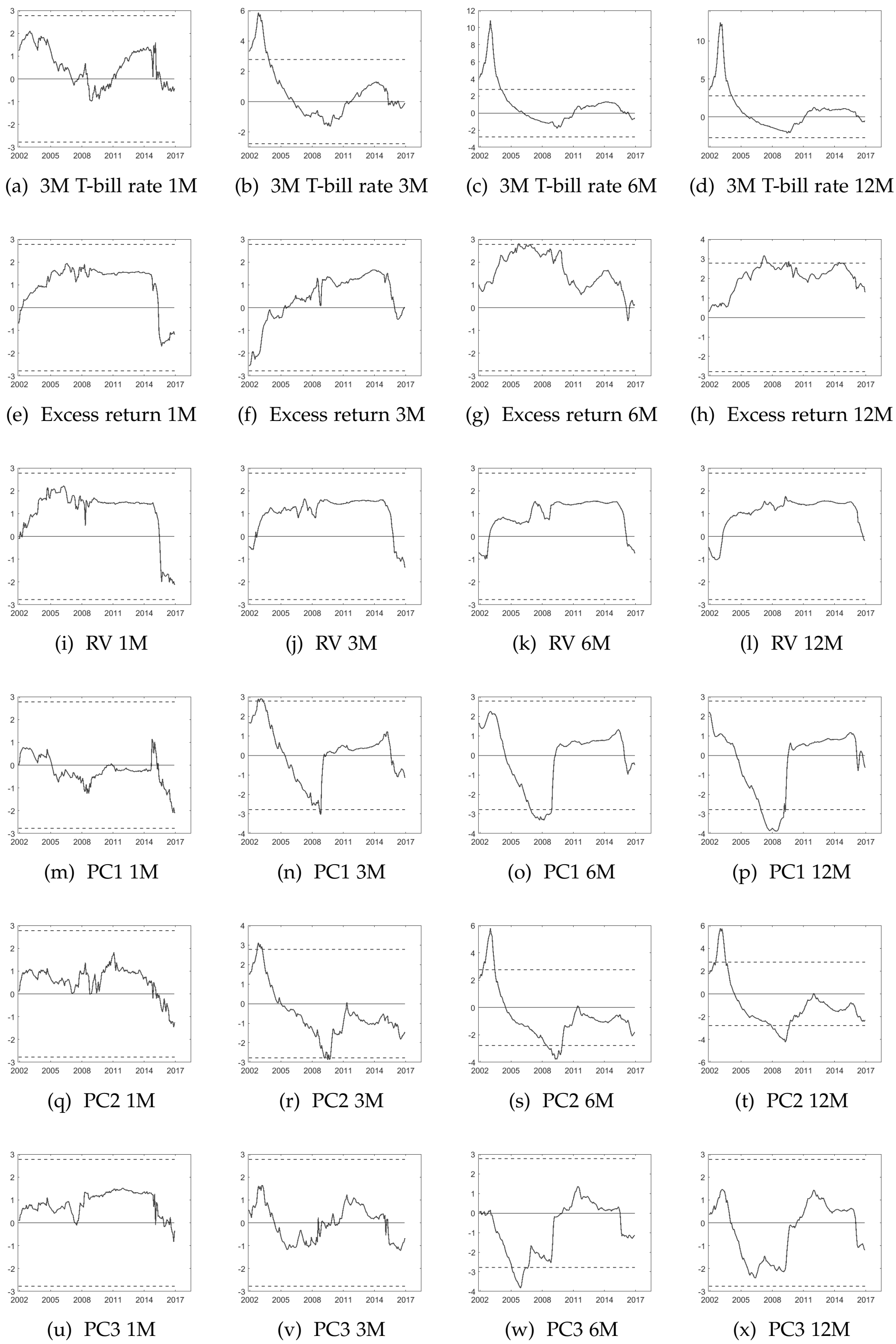
(i) Housing starts 1M (j) Housing starts 3M (k) Housing starts 6M (l) Housing starts 12M

(m) ADS 1M (n) ADS 3M (o) ADS 6M (p) ADS 12M

(q) Term spread 1M (r) Term spread 3M (s) Term spread 6M (t) Term spread 12M

(u) Default spread 1M (v) Default spread 3M (w) Default spread 6M (x) Default spread 12M

Figure 3.18: Fluctuation test results for loss function differences between the GARCH-MIDAS models and the GJR-GARCH(1,1) model. Absolute forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



(a) 3M T-bill rate 1M (b) 3M T-bill rate 3M (c) 3M T-bill rate 6M (d) 3M T-bill rate 12M

(e) Excess return 1M (f) Excess return 3M (g) Excess return 6M (h) Excess return 12M

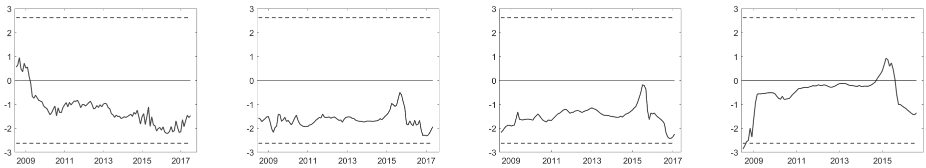
(i) RV 1M (j) RV 3M (k) RV 6M (l) RV 12M

(m) PC1 1M (n) PC1 3M (o) PC1 6M (p) PC1 12M

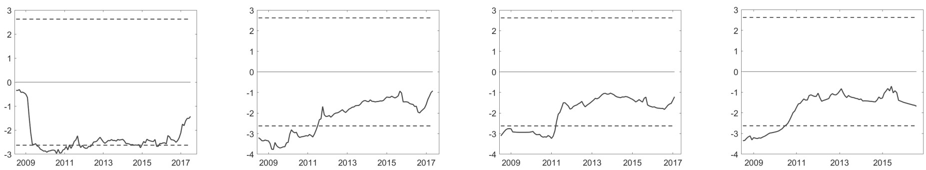
(q) PC2 1M (r) PC2 3M (s) PC2 6M (t) PC2 12M

(u) PC3 1M (v) PC3 3M (w) PC3 6M (x) PC3 12M

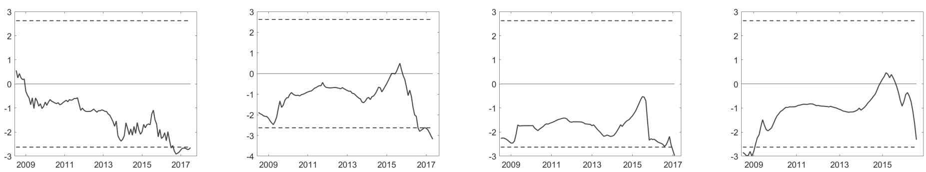
Figure 3.19: Fluctuation test results for loss function differences between the GARCH-MIDAS models and the GJR-GARCH(1,1) model. Absolute forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



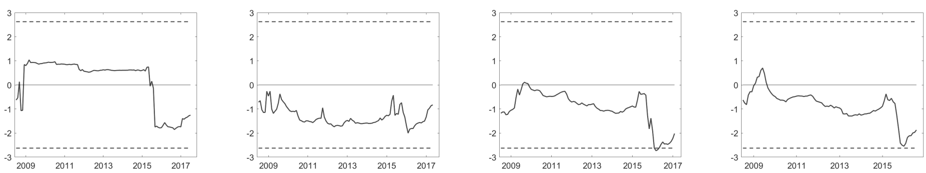
(a) Buying Cond. 1M (b) Buying Cond. 3M (c) Buying Cond. 6M (d) Buying Cond. 12M



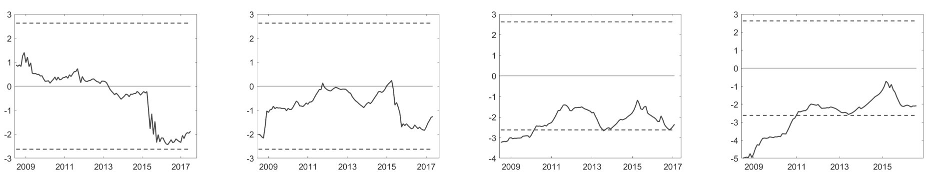
(e) ISM NO 1M (f) ISM NO 3M (g) ISM NO 6M (h) ISM NO 12M



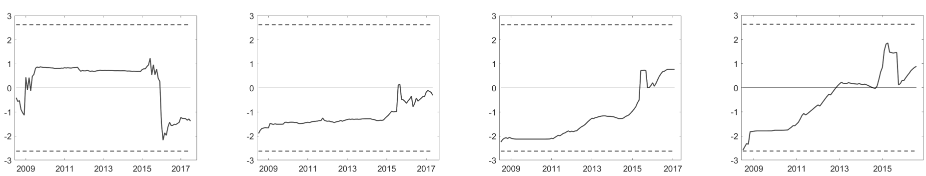
(i) Housing starts 1M (j) Housing starts 3M (k) Housing starts 6M (l) Housing starts 12M



(m) ADS index 1M (n) ADS index 3M (o) ADS index 6M (p) ADS index 12M

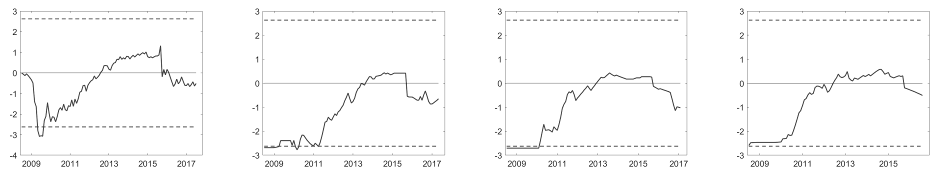


(q) Term spread 1M (r) Term spread 3M (s) Term spread 6M (t) Term spread 12M

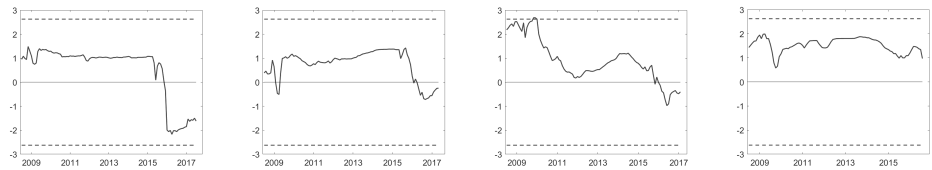


(u) Default spread 1M (v) Default spread 3M (w) Default spread 6M (x) Default spread 12M

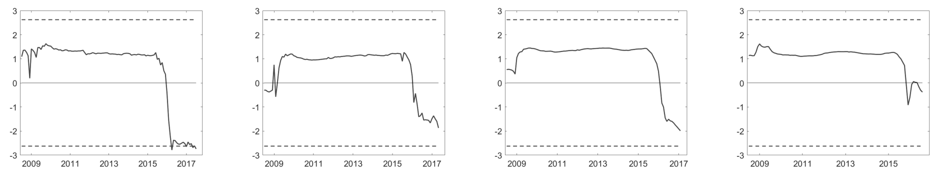
Figure 3.20: Fluctuation test results for loss function differences between the decision rule based forecasts and the GJR-GARCH(1,1) model. Absolute forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



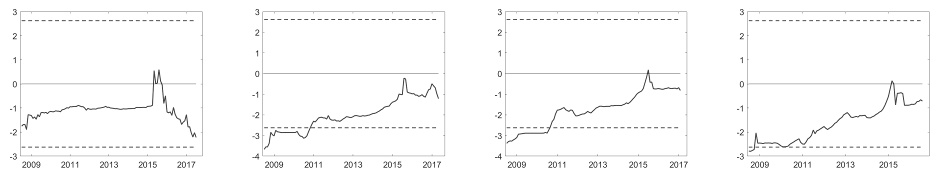
(a) 3M T-bill rate 1M (b) 3M T-bill rate 3M (c) 3M T-bill rate 6M (d) 3M T-bill rate 12M



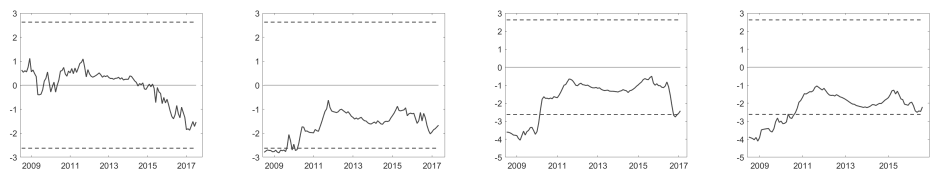
(e) Excess return 1M (f) Excess return 3M (g) Excess return 6M (h) Excess return 12M



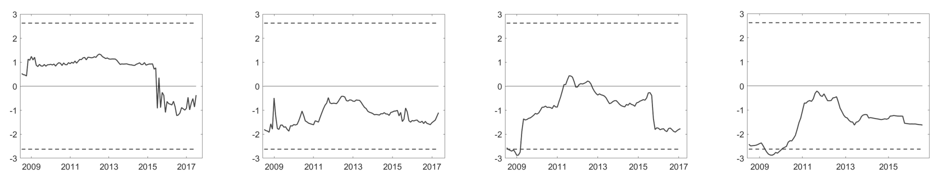
(i) RV 1M (j) RV 3M (k) RV 6M (l) RV 12M



(m) PC1 1M (n) PC1 3M (o) PC1 6M (p) PC1 12M

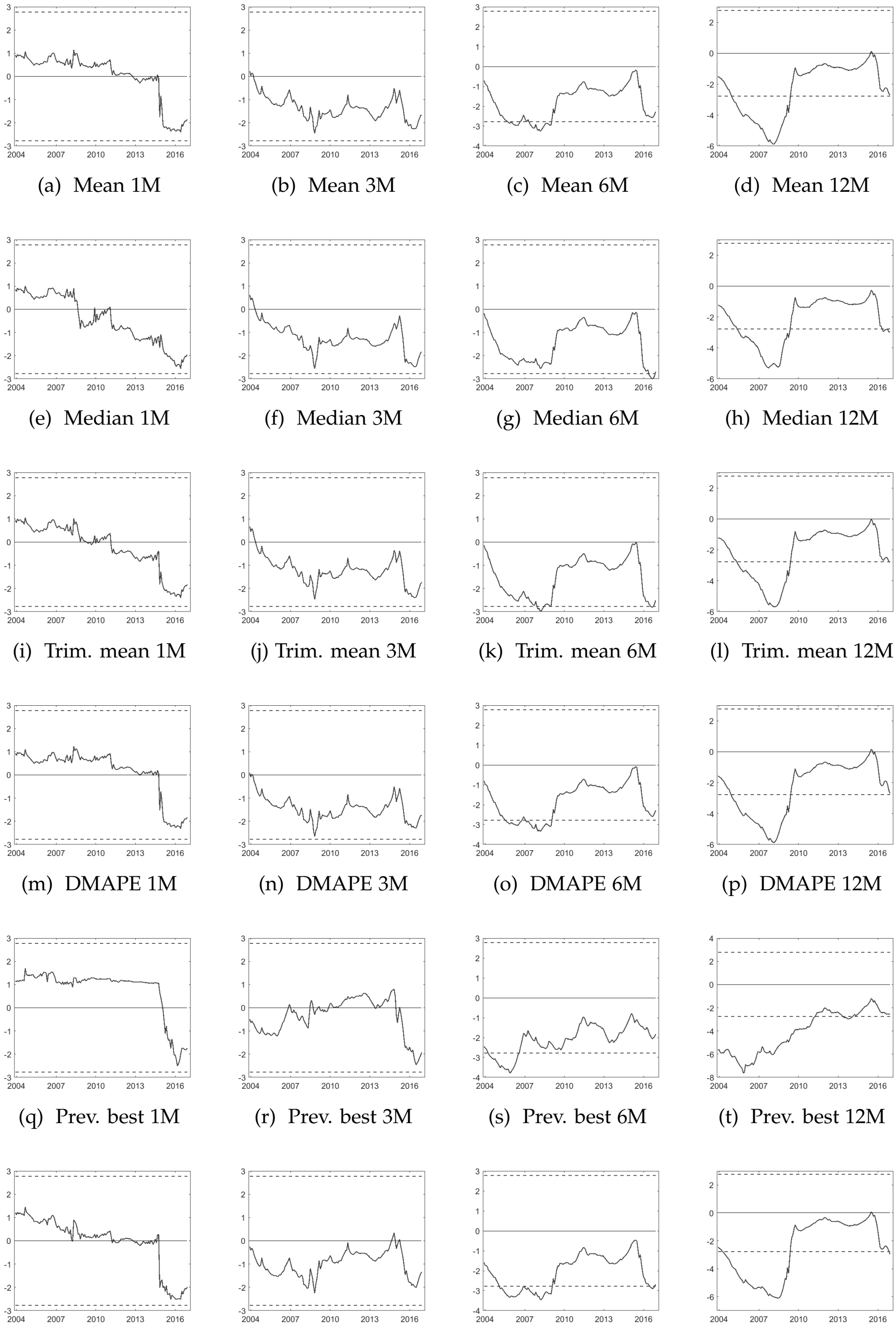


(q) PC2 1M (r) PC2 3M (s) PC2 6M (t) PC2 12M



(u) PC3 1M (v) PC3 3M (w) PC3 6M (x) PC3 12M

Figure 3.21: Fluctuation test results for loss function differences between the decision rule based forecasts and the GJR-GARCH(1,1) model. Absolute forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



(a) Mean 1M (b) Mean 3M (c) Mean 6M (d) Mean 12M

(e) Median 1M (f) Median 3M (g) Median 6M (h) Median 12M

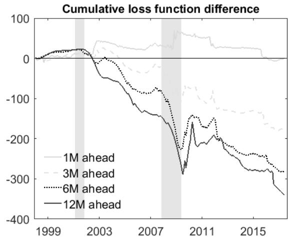
(i) Trim. mean 1M (j) Trim. mean 3M (k) Trim. mean 6M (l) Trim. mean 12M

(m) DMAPE 1M (n) DMAPE 3M (o) DMAPE 6M (p) DMAPE 12M

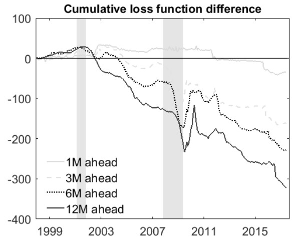
(q) Prev. best 1M (r) Prev. best 3M (s) Prev. best 6M (t) Prev. best 12M

(u) Mean (best 3) 1M (v) Mean (best 3) 3M (w) Mean (best 3) 6M (x) Mean (best 3) 12M

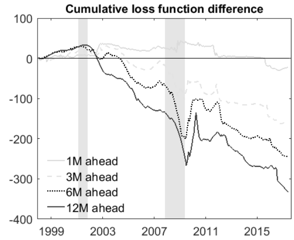
Figure 3.22: Fluctuation test results for loss function differences between the forecast combinations of GARCH-MIDAS models and the GJR-GARCH(1,1) model. Absolute forecast errors. DMAPE with $\eta = 0.5$. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



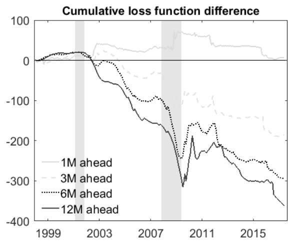
(a) Mean



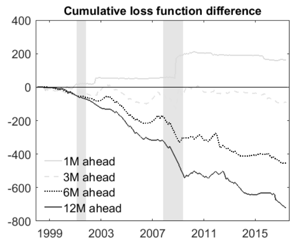
(b) Median



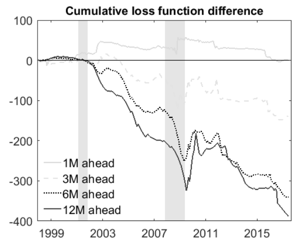
(c) Trimmed mean



(d) DMAPE ($\eta = 0.5$)



(e) Previously best



(f) Mean (best three)

Figure 3.23: Cumulative sum of loss function differences between forecast combinations of GARCH-MIDAS models and the GJR-GARCH(1,1) model ($|Loss_{combo}| - |Loss_{GARCH}|$). An upward sloping segment thus indicates the GJR-GARCH model outperforms the combination forecast. Grey areas mark NBER dated US recessions.

Table 3.11: Effect of business cycle (IP growth) on forecasting performance

	1 month ahead		3 months ahead		6 months ahead		12 months ahead	
	> 0	< 0	> 0	< 0	> 0	< 0	> 0	< 0
Buying Conditions index	1.00	1.00	0.98	0.95*	0.99	0.93*	1.02	0.95***
ISM New Orders index	0.99	0.99	1.00	0.99	1.00	0.98	0.98	0.97*
Housing starts	1.02	0.96	1.03	0.97	1.03	0.94**	1.01	0.90**
ADS index	0.98	1.07	1.01	0.99	1.04	0.96	1.05	0.96*
Term spread	1.01	1.06	0.96	1.01	0.91**	0.92***	0.86***	0.87***
Default spread	1.01	1.16	1.16	1.12	1.39	1.06	1.44	1.04
3M T-bill rate	1.03*	0.99	1.02	1.01	1.03	1.01	1.03	1.00
Excess market return	0.98	1.18*	1.02	1.01	1.07*	1.02	1.16***	1.02
Realised volatility (RV)	1.02	1.25*	1.19	1.24	1.47	1.16	1.64	1.09
Principal component 1	0.99	0.98	1.03	0.99	1.09	0.97	1.08	0.97
Principal component 2	1.02	0.99	0.98	0.97	0.97	0.96**	0.96	0.94**
Principal component 3	1.03	1.08	0.98	1.01	1.00	0.96**	1.02	0.97**

Benchmark: GJR-GARCH(1,1). MAFE ratio: $\frac{MAFE_{GMX}}{MAFE_{GARCH}}$, where $MAFE_{GMX}$ stands for the mean absolute forecast error of the GARCH-MIDAS-X model. A value below 1 means the GARCH-MIDAS-X model outperforms the GJR-GARCH(1,1) model. Otherwise, see notes on Table 3.4.

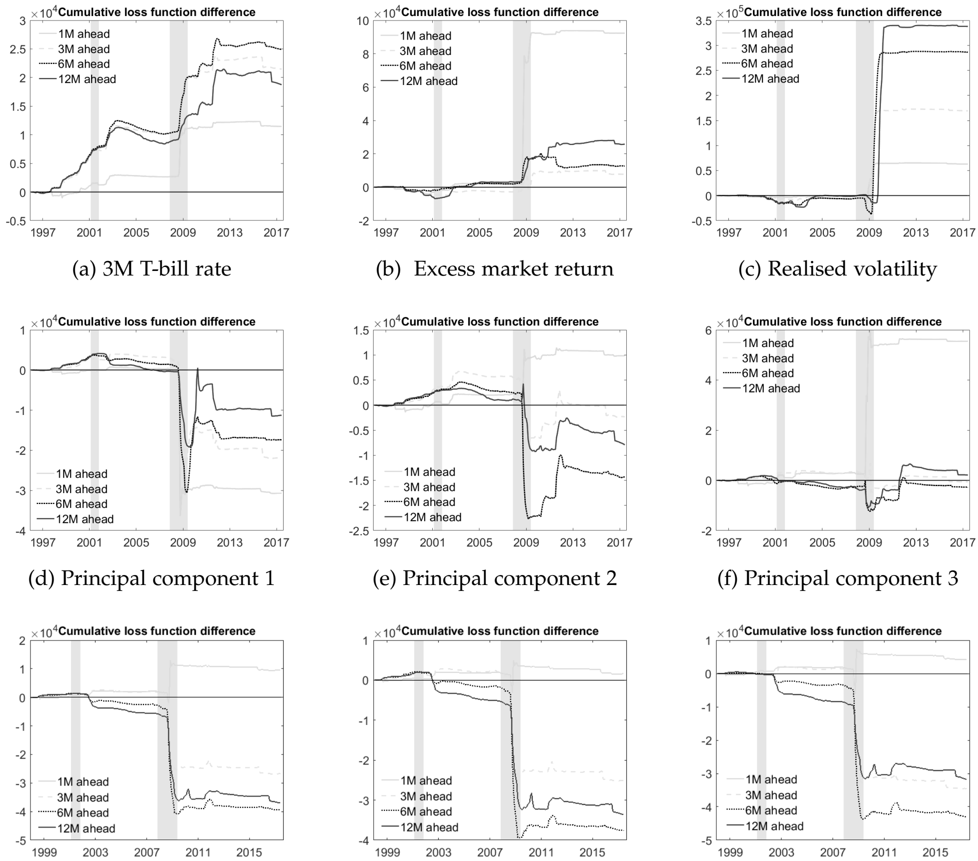
Table 3.12: Effect of volatility environment on forecasting performance

	1 month ahead		3 months ahead		6 months ahead		12 months ahead	
	Low	High	Low	High	Low	High	Low	High
Buying Conditions index	0.99	1.01	0.89**	0.98	0.83***	0.99	0.84**	1.01
ISM New Orders index	0.97	0.99	0.88**	1.02	0.84***	1.02	0.79***	1.02
Housing starts	0.97	0.99	0.91*	1.02	0.88**	1.00	0.84*	0.97
ADS index	0.99	1.03	0.96	1.01	0.94*	1.01	0.92*	1.02
Term spread	1.01	1.04	0.85*	1.02	0.71***	0.96	0.58***	0.94**
Default spread	0.99	1.11	0.97	1.18	1.00	1.25	1.14	1.22
3M T-bill rate	0.95***	1.02	0.84***	1.05***	0.80***	1.07***	0.76***	1.07***
Excess market return	0.97	1.11*	0.99	1.02	1.13**	1.02	1.20**	1.05*
Realised volatility (RV)	0.94	1.19**	0.83**	1.31**	0.89	1.38	1.22	1.33
Principal component 1	0.97	0.99	0.96	1.02	0.97	1.04	0.97	1.03
Principal component 2	0.94**	1.02	0.78***	1.02	0.68***	1.03	0.61***	1.03*
Principal component 3	0.98	1.07	0.83***	1.03*	0.76***	1.03	0.70***	1.06**

*Benchmark: GJR-GARCH(1,1) model. MAFE ratio: $\frac{MAFE_{GMX}}{MAFE_{GARCH}}$, where $MAFE_{GMX}$ stands for the mean absolute forecast error of the GARCH-MIDAS-X model. A value below 1 means the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test. High / low volatility months are classified according to the median of the VIX index: 147 months of high volatility and 111 months of low volatility. $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$.*

3.E Additional results on time-varying forecasting performance

This appendix presents the remaining cumulative sums of the loss function differences and the results of the Fluctuation test, complementing the results in Section 3.6.2. The decision to exclude these figures from the main text relies on three considerations. First, the Fluctuation test results for the 3M T-bill rate are mostly similar to those of the default spread. Second, the excess market return and realised volatility lead to a generally weak performance throughout the sample period, as was already clear from the full-sample results, making the time-varying results less interesting. Third, the principal components driven models lead to largely similar, and at least no better, forecast accuracy than the series they are based on.

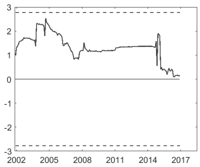


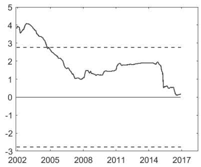
(a) 3M T-bill rate (b) Excess market return (c) Realised volatility

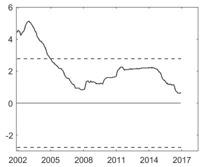
(d) Principal component 1 (e) Principal component 2 (f) Principal component 3

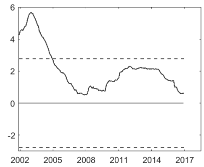
(g) Mean (h) Trimmed mean (i) Mean (best three)

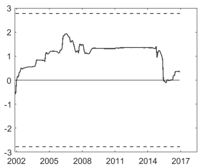
Figure 3.24: Cumulative sum of loss function differences between forecast combinations of GARCH-MIDAS models and the GJR-GARCH(1,1) model ($Loss^2_{combo} - Loss^2_{GARCH}$). An upward sloping segment thus indicates the GJR-GARCH model outperforms the GARCH-MIDAS model. Grey areas mark NBER dated US recessions.

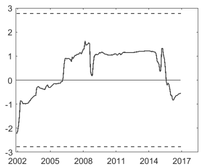

(a) 3M T-bill rate 1M

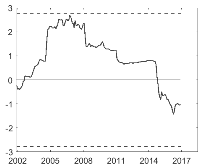

(b) 3M T-bill rate 3M

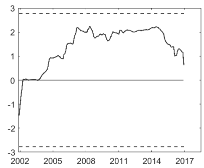

(c) 3M T-bill rate 6M

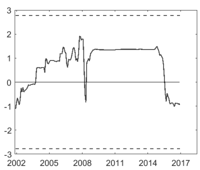

(d) 3M T-bill rate 12M

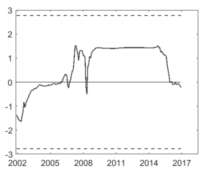

(e) Excess return 1M

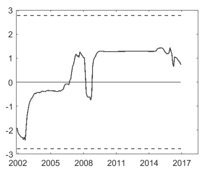

(f) Excess return 3M

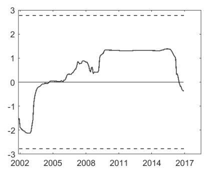

(g) Excess return 6M

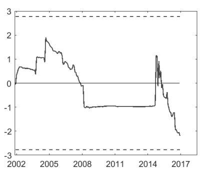

(h) Excess return 12M

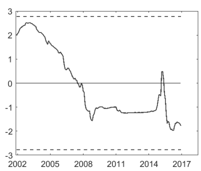

(i) RV 1M

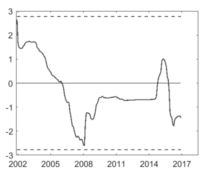

(j) RV 3M

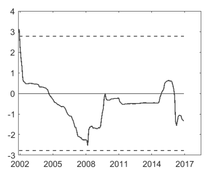

(k) RV 6M

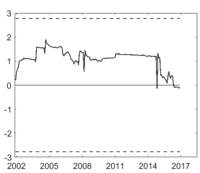

(l) RV 12M

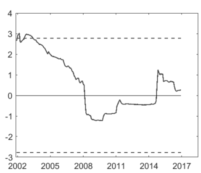

(m) PC1 1M

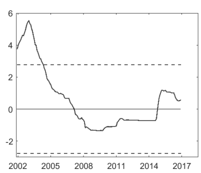

(n) PC1 3M

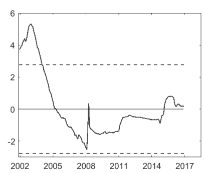

(o) PC1 6M

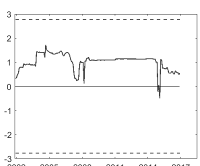

(p) PC1 12M

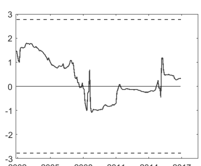

(q) PC2 1M

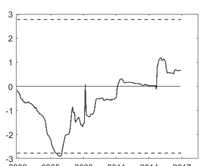

(r) PC2 3M


(s) PC2 6M


(t) PC2 12M


(u) PC3 1M


(v) PC3 3M


(w) PC3 6M

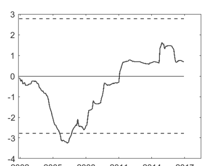
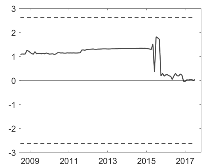
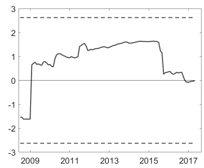

(x) PC3 12M

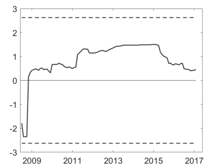
Figure 3.25: Fluctuation test results for loss function differences between the GARCH-MIDAS model and the GJR-GARCH(1,1) model. Squared forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



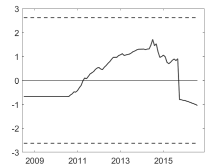
(a) 3M T-bill rate 1M



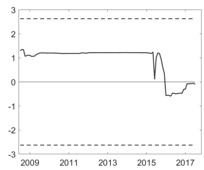
(b) 3M T-bill rate 3M



(c) 3M T-bill rate 6M

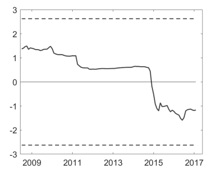


(d) 3M T-bill rate 12M

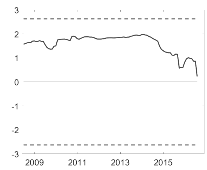


(e) Excess return 1M

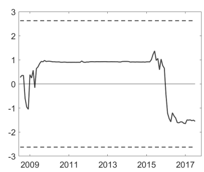
(f) Excess return 3M



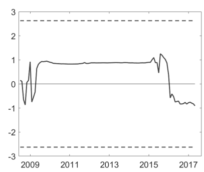
(g) Excess return 6M



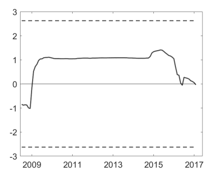
(h) Excess return 12M



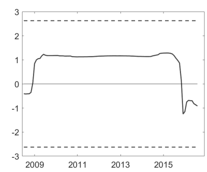
(i) RV 1M



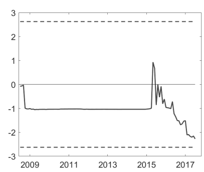
(j) RV 3M



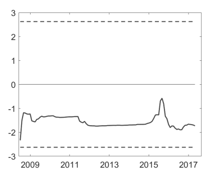
(k) RV 6M



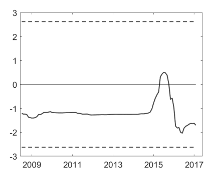
(l) RV 12M



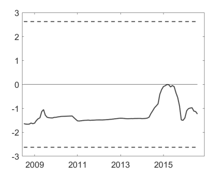
(m) PC1 1M



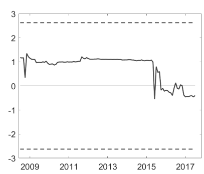
(n) PC1 3M



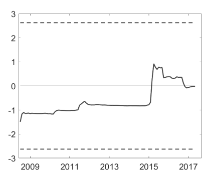
(o) PC1 6M



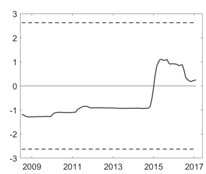
(p) PC1 12M



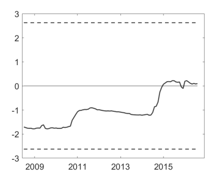
(q) PC2 1M



(r) PC2 3M

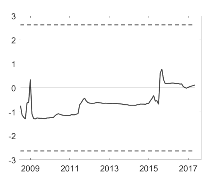


(s) PC2 6M

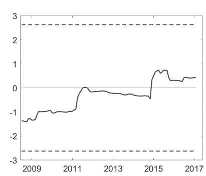


(t) PC2 12M

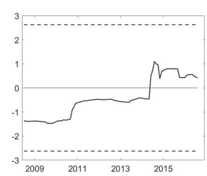
(u) PC3 1M



(v) PC3 3M

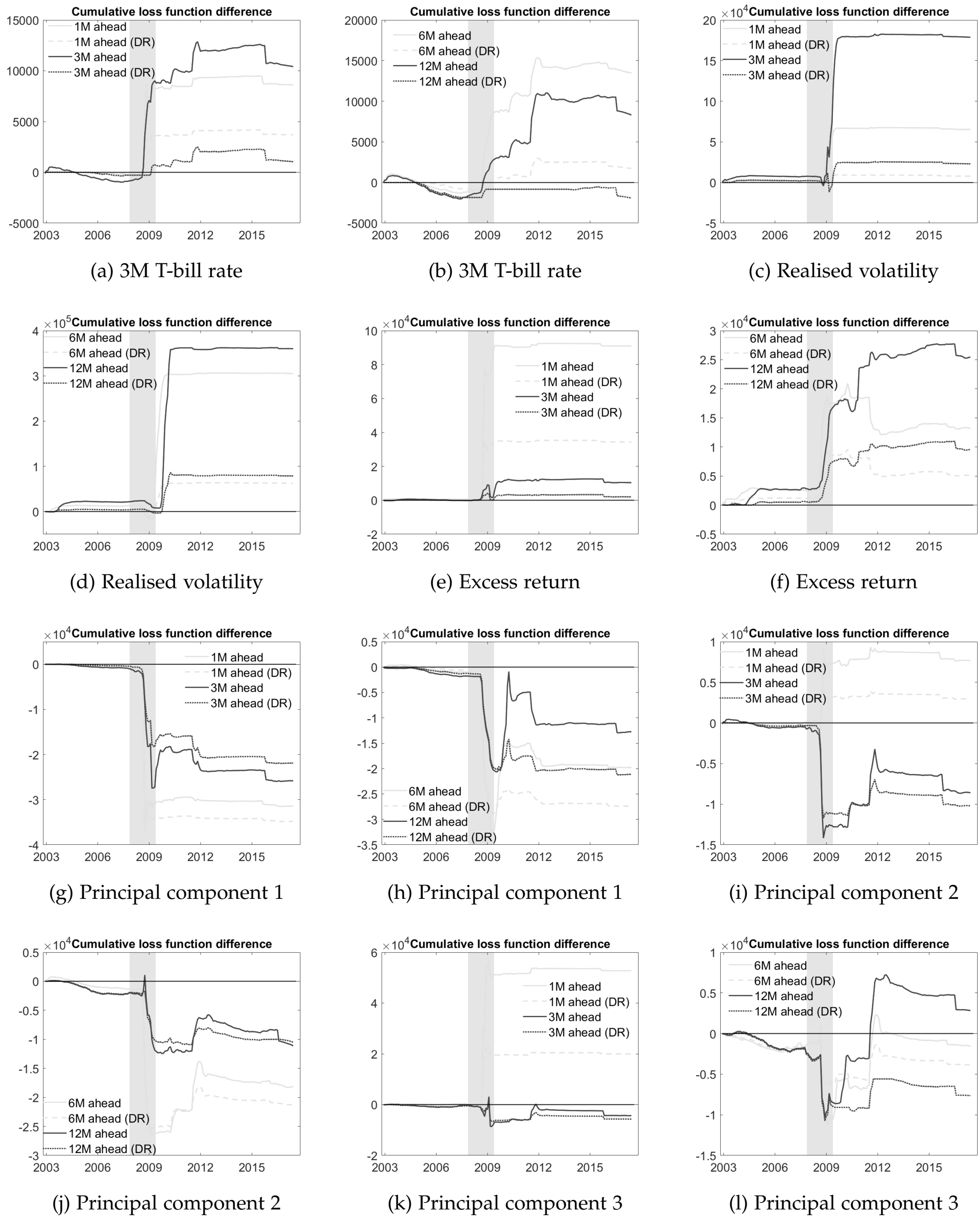


(w) PC3 6M



(x) PC3 12M

Figure 3.26: Fluctuation test results for loss function differences between the decision rule based forecast and the GJR-GARCH(1,1) model. Squared forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.



(a) 3M T-bill rate

(b) 3M T-bill rate

(c) Realised volatility

(d) Realised volatility

(e) Excess return

(f) Excess return

(g) Principal component 1

(h) Principal component 1

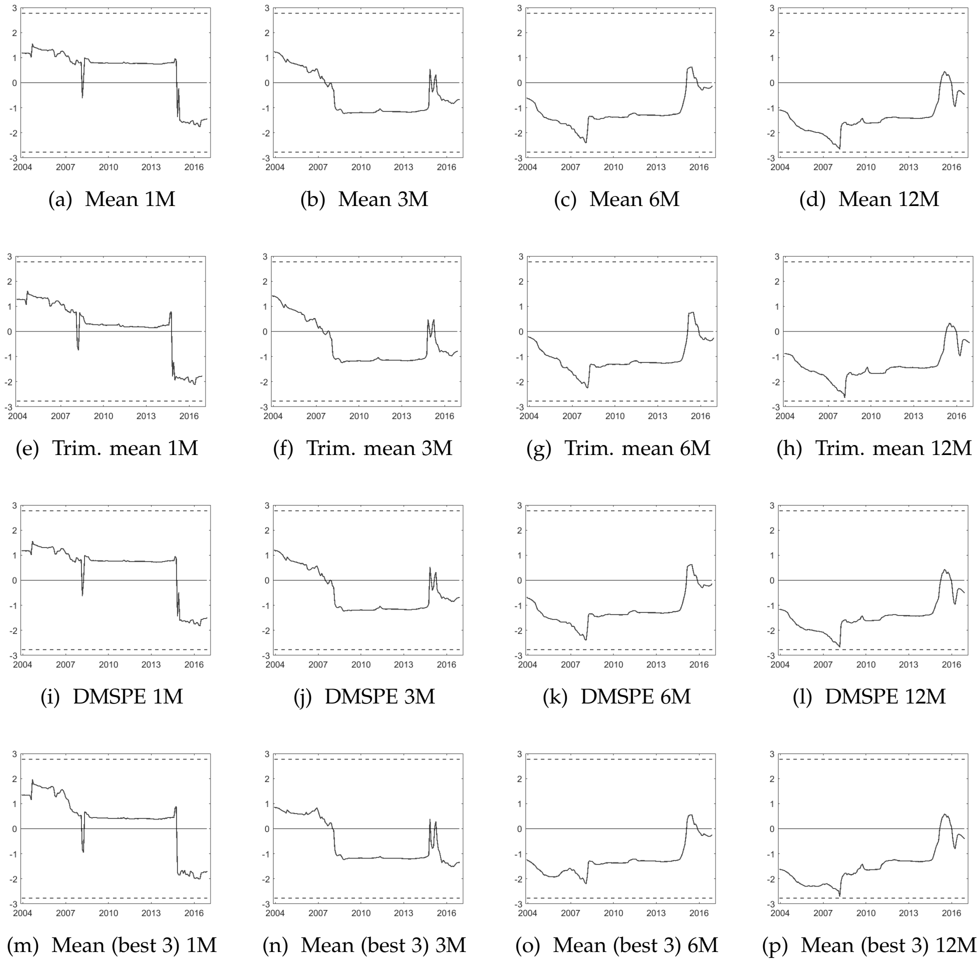
(i) Principal component 2

(j) Principal component 2

(k) Principal component 3

(l) Principal component 3

Figure 3.27: Cumulative sum of loss function differences between the decision rule (DR) based forecasts and the GJR-GARCH(1,1) model ($Loss^2_{DR} - Loss^2_{GARCH}$). An upward sloping segment thus indicates the GJR-GARCH model outperforms the decision rule. Grey areas mark NBER dated US recessions.



(a) Mean 1M (b) Mean 3M (c) Mean 6M (d) Mean 12M

(e) Trim. mean 1M (f) Trim. mean 3M (g) Trim. mean 6M (h) Trim. mean 12M

(i) DMSPE 1M (j) DMSPE 3M (k) DMSPE 6M (l) DMSPE 12M

(m) Mean (best 3) 1M (n) Mean (best 3) 3M (o) Mean (best 3) 6M (p) Mean (best 3) 12M

Figure 3.28: Fluctuation test results for loss function differences between forecast combinations of GARCH-MIDAS models and the GJR-GARCH(1,1) model. Dashed lines represent 90% confidence bands. DMSPE with $\eta = 1$. The year on the x-axis marks the end of the rolling window period over which the test statistics is calculated.

3.F Robustness of estimation and weighting schemes

In this appendix I discuss the robustness of the results to (i) the weighting scheme, i.e., fixed weights instead of weights re-estimated each period, and (ii) the estimation scheme, i.e., using an expanding window instead of a rolling window. The first part thus considers whether there arises any instability from re-estimating the weights in each period and whether this instability impacts the forecasts, and if yes, in which way. The second part considers whether the flexibility we gain from using a rolling window leads to more or less accurate in-sample fit and out-of-sample forecasts compared to an expanding window estimation scheme, which is less prone to instability but also less flexible. Thus I have, first of all, estimated the models over the full sample, saved the weights of the weighting schemes and then re-estimated the models using a rolling window with the weights fixed at the full-sample weights. The other parameters of the GARCH-MIDAS model are re-estimated each period. Secondly, I have estimated each GARCH-MIDAS model using an expanding window, i.e., adding one month to the estimation window in each period. The differences in the forecasts produced, the in-sample fit (in terms of the variance ratio) and the parameter estimates are discussed below.

Table 3.13: Forecast comparison: fixed vs. re-estimated weights

	1M ahead		3M ahead		6M ahead		12M ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Buying Conditions index	0.99	0.99	0.99	1.02	1.00	1.01	1.00	1.00
ISM New Orders index	1.00	0.99	0.99	0.99	1.01	0.97	1.01	1.00
Housing starts	1.01	1.01	0.99	1.00	1.00	1.02	1.03	1.05
ADS index	0.97	0.80	1.02	1.03	1.04	1.03	1.02	1.02
Term spread	1.00	0.99	0.99	1.00	1.00	1.00	1.01	1.00
Default spread	1.00	1.05	0.99	0.98	1.00	1.00	1.00	1.01
3M T-bill rate	1.00	1.00	0.99	1.00	0.99	1.00	0.99	1.00
Excess market return	0.93	0.81	0.98	0.98	0.96	0.98	0.92	0.96
Realised volatility (RV)	1.00	1.11	1.01	1.08	1.01	1.11	1.00	1.12
First PC	1.00	1.01	1.01	1.01	1.01	1.01	1.01	1.01
Second PC	0.99	0.99	1.00	1.00	1.00	1.01	1.00	0.99
Third PC	0.95	0.83	0.98	1.01	0.98	1.00	0.95	0.98

MAFE ratio: $\frac{MAFE_{GMXfix}}{MAFE_{GMX}}$*, where* $MAFE_{GMXfix}$ *(*$MAFE_{GMX}$*) stands for the mean absolute forecast error of the GARCH-MIDAS-X model estimated using fixed (re-estimated) weights. The MSFE ratio is calculated equivalently. A value below 1 means the forecast based on fixed weights outperforms the forecast based on weights re-estimated each period.* $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$.

Starting with (i), over the full sample the choice of fixed or rolling weights has little effect on the out-of-sample forecasts (Table 3.13). Fixing the weights improves forecasting performance consistently and relatively clearly only for the excess market return and third principal component. Figures 3.30 and 3.31 look at the cumulative loss function differences vis-à-vis the GJR-GARCH(1,1) model. By comparing the solid line with the dashed line it can be seen that in most cases the weighting scheme does not matter much for the relative performance of the models over time. However, for the ADS index and housing starts fixing the weight(s) clearly matter during/after the financial crisis: the 12M (and also 3M for the ADS index) ahead forecasts perform worse when fixing the weights, while for the ADS index the 1M ahead forecast performs clearly better. For excess returns and PC3 the forecasts using fixed weights perform clearly better in most periods than the forecasts from the models where the weights are re-estimated each period, confirming the full-sample results.

Mostly the differences in in-sample fit (variance ratios) are relatively small (Figure 3.32, compare the dark blue and light blue lines). The exception is the GARCH-MIDAS model driven by the term spread towards the end of the period, the PC3 driven model in the early part of the sample, and the model driven by excess market returns. In all cases the model with weight parameters re-estimated each period produces a better fit.

Figure 3.33 shows the estimates for θ for fixed and re-estimated weights (compare the dark blue and light blue lines). The differences are mostly small. The main exception is again excess market returns, for which θ has the opposite sign when weights are fixed, compared to other estimation schemes. However, as we can see from Figure 3.34 the negative θ estimate from the fixed weights model is only borderline statistically significant. The change in sign of the second and third PC still occurs regardless of the weighting scheme, suggesting it is not a consequence of imprecisely estimated weights but rather the changing composition of the PC.

Table 3.14: Forecast comparison: expanding vs. rolling window

	1M ahead		3M ahead		6M ahead		12M ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
GJR-GARCH(1,1)	0.98	1.03	0.97	0.98	0.95	0.99	0.97	1.00
Buying Conditions index	0.98	1.09	0.99	1.08	0.98	1.06	0.96	1.01
ISM New Orders index	0.98	1.04	0.96	0.99	0.96	1.00	0.97	1.00
Housing starts	0.99	1.09	0.95	1.02	0.96	1.04	0.98	1.04
ADS index	0.95	0.87	0.97	1.03	0.98	1.02	0.98	1.01
Term spread	0.97	0.98	0.97	1.02	1.00	1.04	1.01	1.04
Default spread	0.90	0.87	0.85	0.88	0.82	0.83	0.83	0.81
3M T-bill rate	0.97	1.02	0.95	0.98	0.95	0.99	0.96	1.00
Excess market return	0.96	0.90	0.96	0.97	0.94	0.98	0.94	0.97
Realised volatility (RV)	0.96	0.98	0.95	0.96	0.96	0.96	0.96	0.96
First PC	0.99	1.12	0.97	1.01	0.96	1.01	0.96	1.00
Second PC	0.97	1.05	0.98	1.02	0.99	1.04	1.00	1.02
Third PC	0.94	0.89	0.96	1.02	0.96	1.01	0.95	1.00

MAFE ratio: $\frac{MAFE_{GMXexp}}{MAFE_{GMX}}$*, where* $MAFE_{GMXexp}$ *(*$MAFE_{GMX}$*) stands for the mean absolute forecast error of the GARCH-MIDAS-X model estimated using an expanding (rolling) window estimation scheme. The MSFE ratio is calculated equivalently. A value below 1 means the expanding window forecast outperforms the rolling window forecast.* $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$.

Moving on to (ii), the expanding window estimation scheme leads to lower fore-

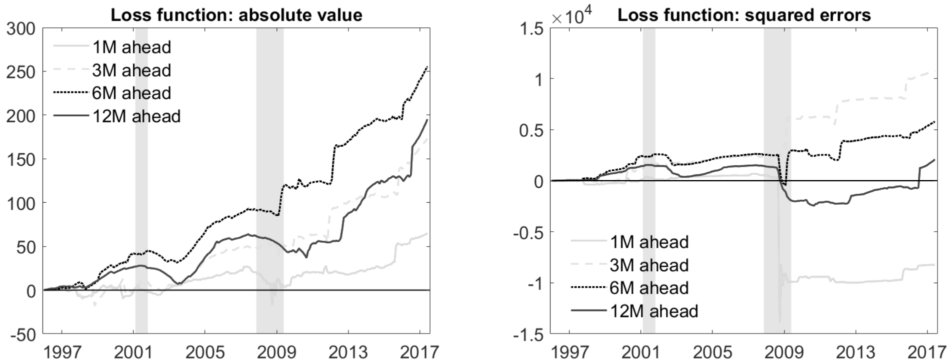
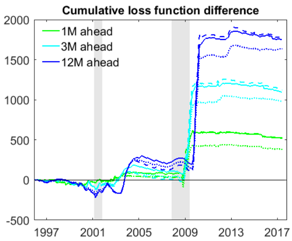


Figure 3.29: Cumulative sum of loss function differences of the GJR-GARCH model estimated using a rolling window and an expanding window. An upward sloping segment indicates the model estimated using the expanding window outperforms the model estimated using a rolling window.

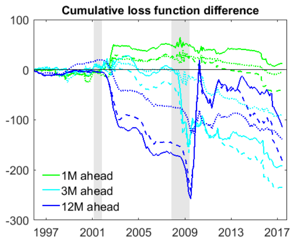
cast errors when using the absolute value of the forecast error as the loss function (Table 3.14). When looking at the MSFE ratios the expanding window leads to lower forecast errors primarily for the financial data. For macroeconomic data and the principal components the rolling window scheme generally leads to lower forecast errors, implying the added flexibility of the rolling window is beneficial.

As expected, the expanding window estimation scheme leads to more stable parameter estimates, which are closer to the full-sample estimates for all models (Figures 3.33 and 3.35). If we compare these results to those in Table 3.14 it is clear that, for example, for the default spread the increased flexibility of the rolling window is not beneficial from a forecasting perspective, whereas for the Buying Conditions index and housing starts it can be. Figure 3.34 indicates that the estimate of θ also tend to be more strongly statistically significant for the expanding window estimation scheme, owing potentially to the larger sample size they are estimated on. From Figure 3.32 we can see that also the variance ratios fluctuate less when an expanding scheme is used, but they tend to be lower, implying a worse in-sample fit.

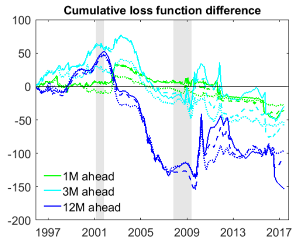
Regarding the cumulative sum of loss function differences, there is a significant difference already for the benchmark GJR-GARCH model (Figure 3.29) as the expanding window scheme performs better in most periods. When comparing the losses of the GARCH-MIDAS models to those of the GJR-GARCH model (Figures 3.30 and 3.31, compare solid line with dotted line) we see that large differences, in favour of the expanding window scheme, occur for, for example, excess returns (all horizons), the ADS index (1M horizon), the default spread (all horizons) and the 3M T-bill rate (1M and 3M horizon). The rolling window estimation scheme leads to more accurate forecasts (relative to the benchmark) when the GARCH-MIDAS model is driven by, for example, the term spread (3M and 12M horizons), housing starts (all horizons), the Buying Conditions index (3M and 12M horizons) and the second principal component (all horizons). The differences in forecasting performance mostly arise in conjunction with the financial crisis.



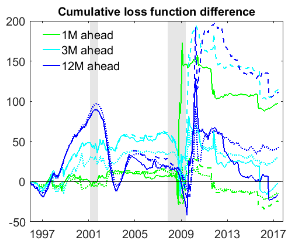
(a) Realised volatility



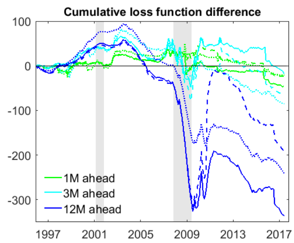
(b) Buying Conditions



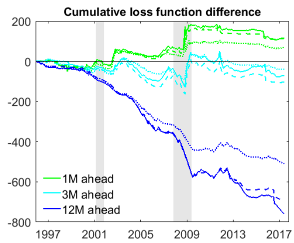
(c) ISM New Orders



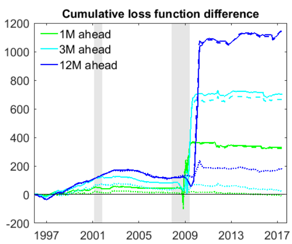
(d) ADS index



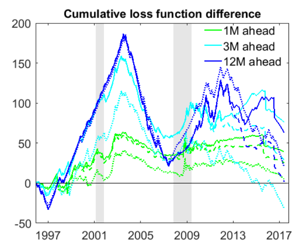
(e) Housing starts



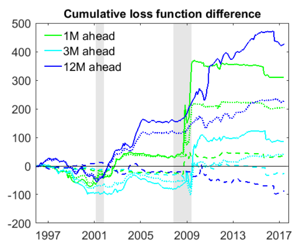
(f) Term spread



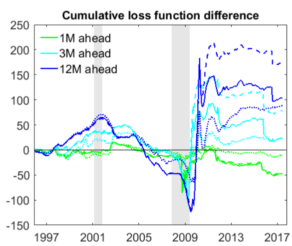
(g) Default spread



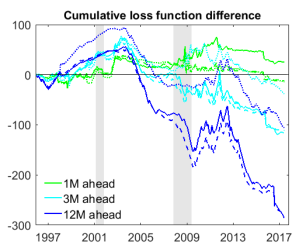
(h) 3M T-bill rate



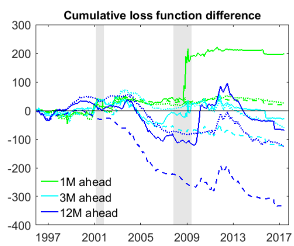
(i) Excess market return



(j) Principal component 1

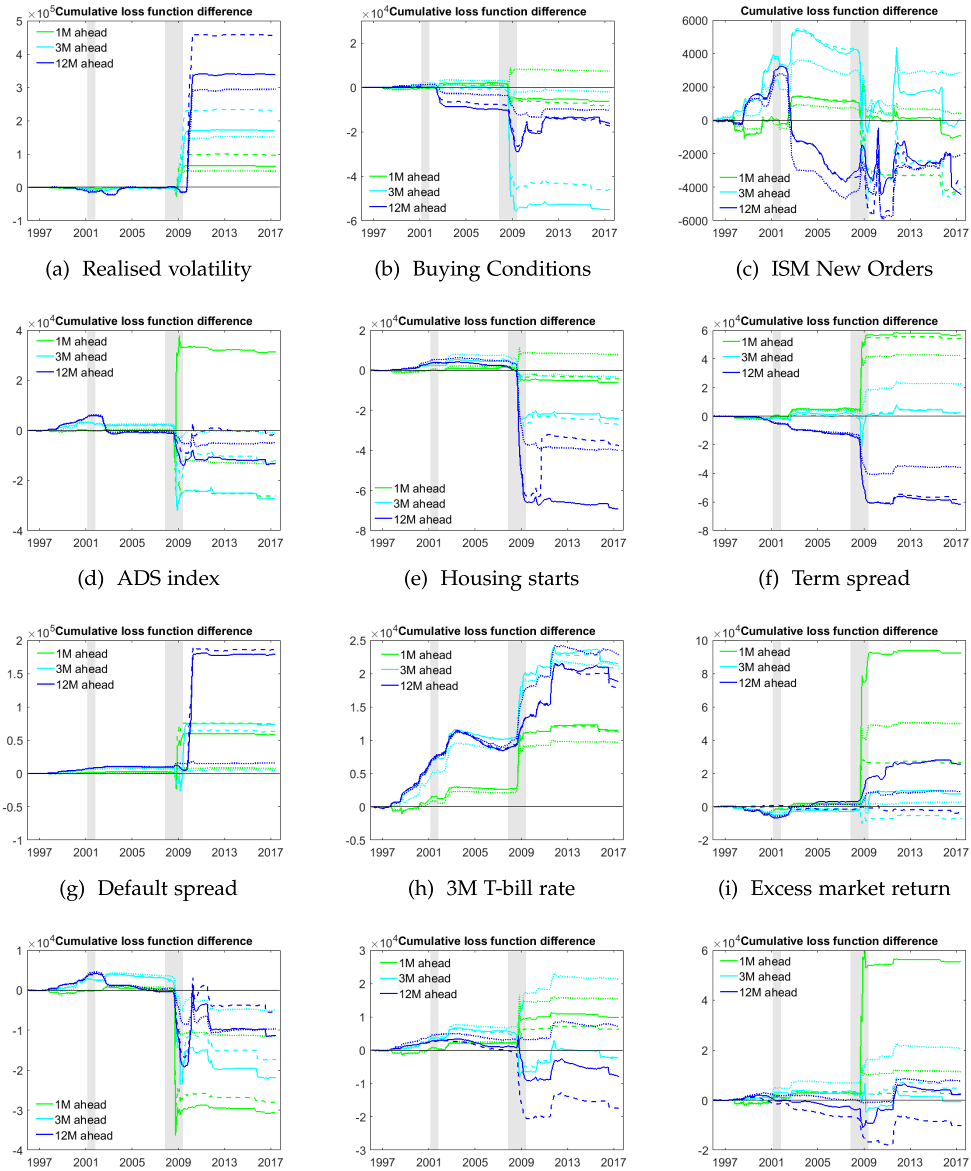


(k) Principal component 2



(l) Principal component 3

Figure 3.30: Cumulative sum of loss function differences (absolute errors) of rolling window GARCH-MIDAS models with fixed weights (dashed line), GARCH-MIDAS models estimated using an expanding window (dotted line), and rolling window GARCH-MIDAS models with weights re-estimated each period (solid line). Baseline model: the GJR-GARCH(1,1) model, estimated using either a rolling window or an expanding window. When the line is upward sloping the GJR-GARCH model outperforms the GARCH-MIDAS model.



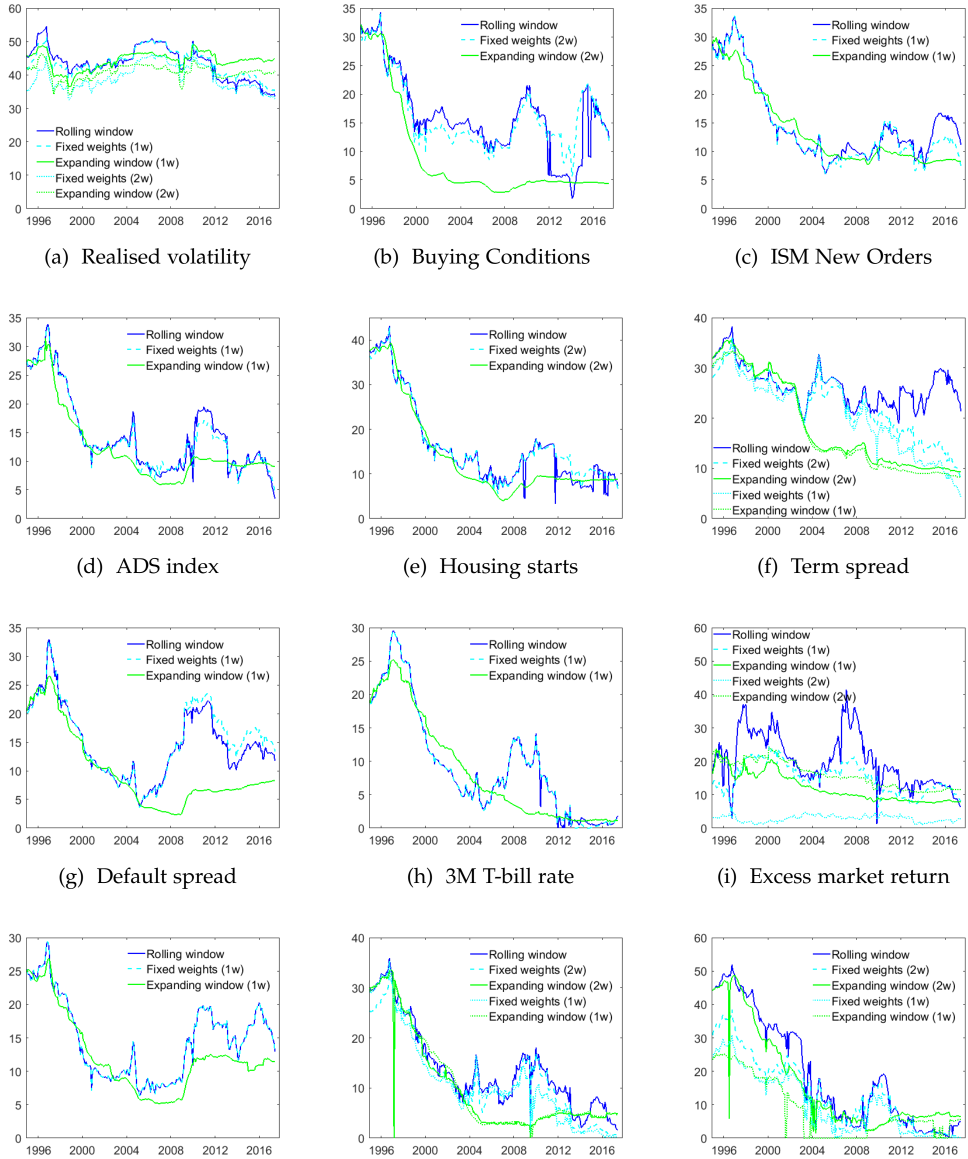
(a) Realised volatility (b) Buying Conditions (c) ISM New Orders

(d) ADS index (e) Housing starts (f) Term spread

(g) Default spread (h) 3M T-bill rate (i) Excess market return

(j) Principal component 1 (k) Principal component 2 (l) Principal component 3

Figure 3.31: Cumulative sum of loss function differences (squared errors) of rolling window GARCH-MIDAS models with fixed weights (dashed line), GARCH-MIDAS models estimated using an expanding window (dotted line), and rolling window GARCH-MIDAS models with weights re-estimated each period (solid line). Baseline model: the GJR-GARCH(1,1) model, estimated using either a rolling window or an expanding window. When the line is upward sloping the GJR-GARCH model outperforms the GARCH-MIDAS model.



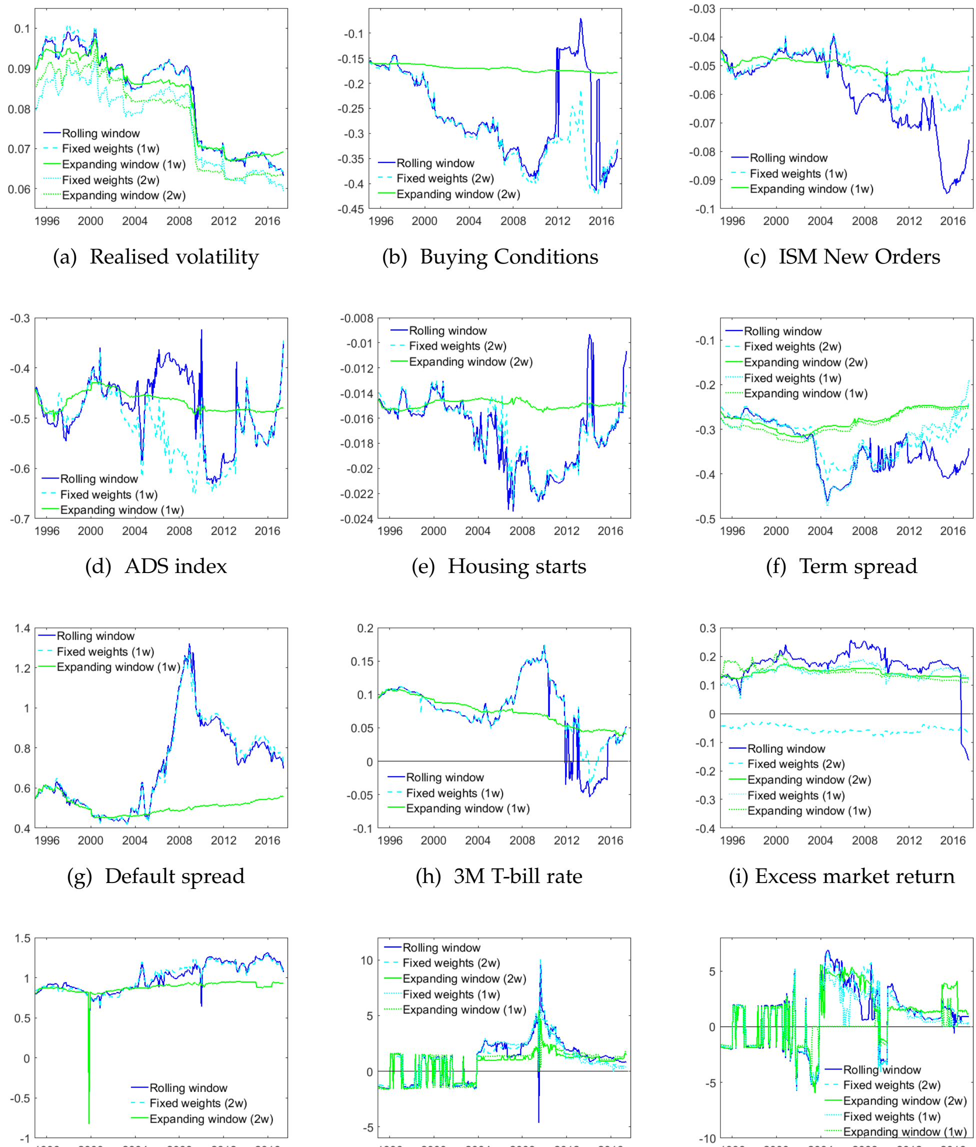
(a) Realised volatility (b) Buying Conditions (c) ISM New Orders

(d) ADS index (e) Housing starts (f) Term spread

(g) Default spread (h) 3M T-bill rate (i) Excess market return

(j) Principal component 1 (k) Principal component 2 (l) Principal component 3

Figure 3.32: Variance ratios of the rolling window GARCH-MIDAS models with fixed weights and the rolling and expanding window GARCH-MIDAS models with the weights re-estimated each period.



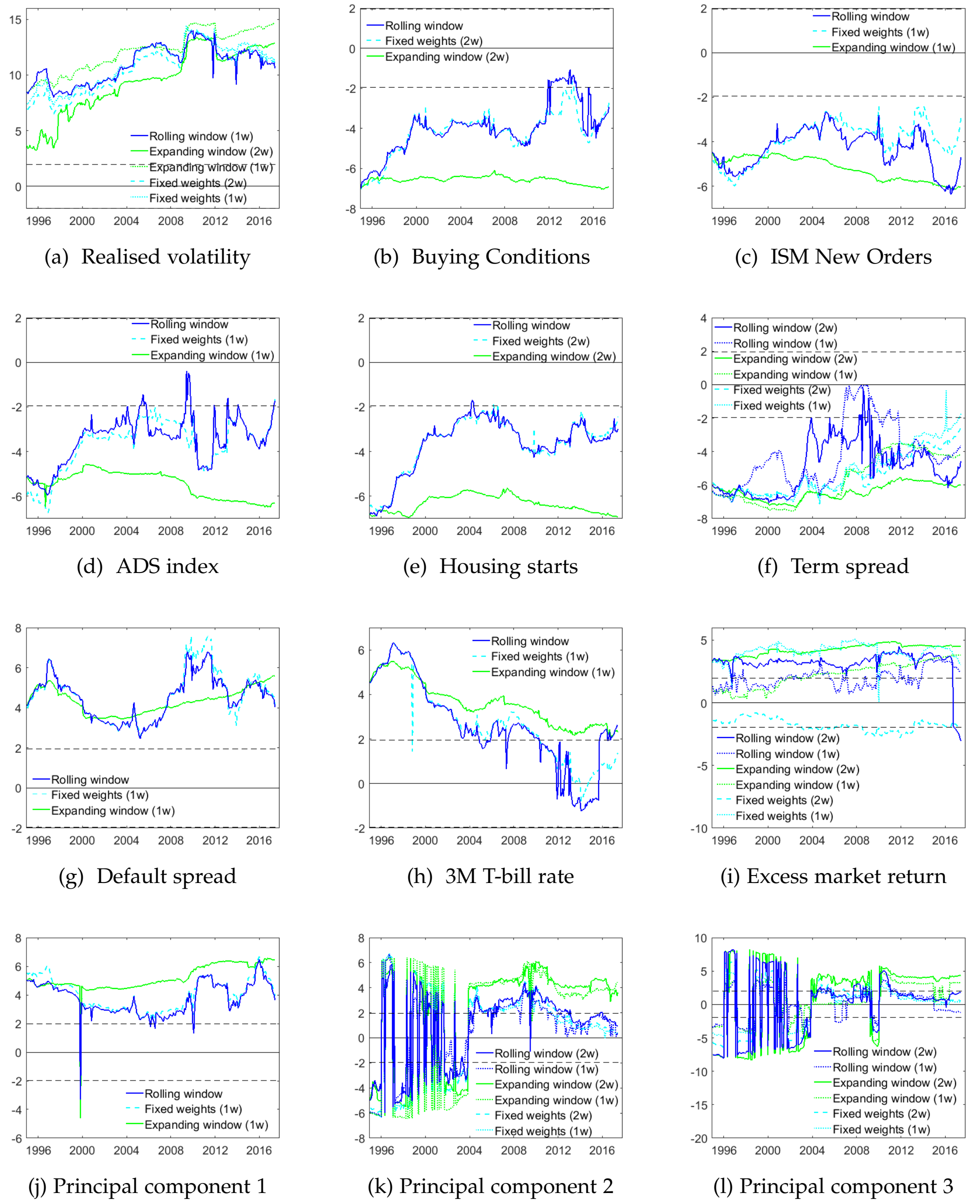
(a) Realised volatility (b) Buying Conditions (c) ISM New Orders

(d) ADS index (e) Housing starts (f) Term spread

(g) Default spread (h) 3M T-bill rate (i) Excess market return

(j) Principal component 1 (k) Principal component 2 (l) Principal component 3

Figure 3.33: Estimates of θ from the rolling window GARCH-MIDAS models with fixed weights and the rolling and expanding window GARCH-MIDAS models with the weights re-estimated each period.



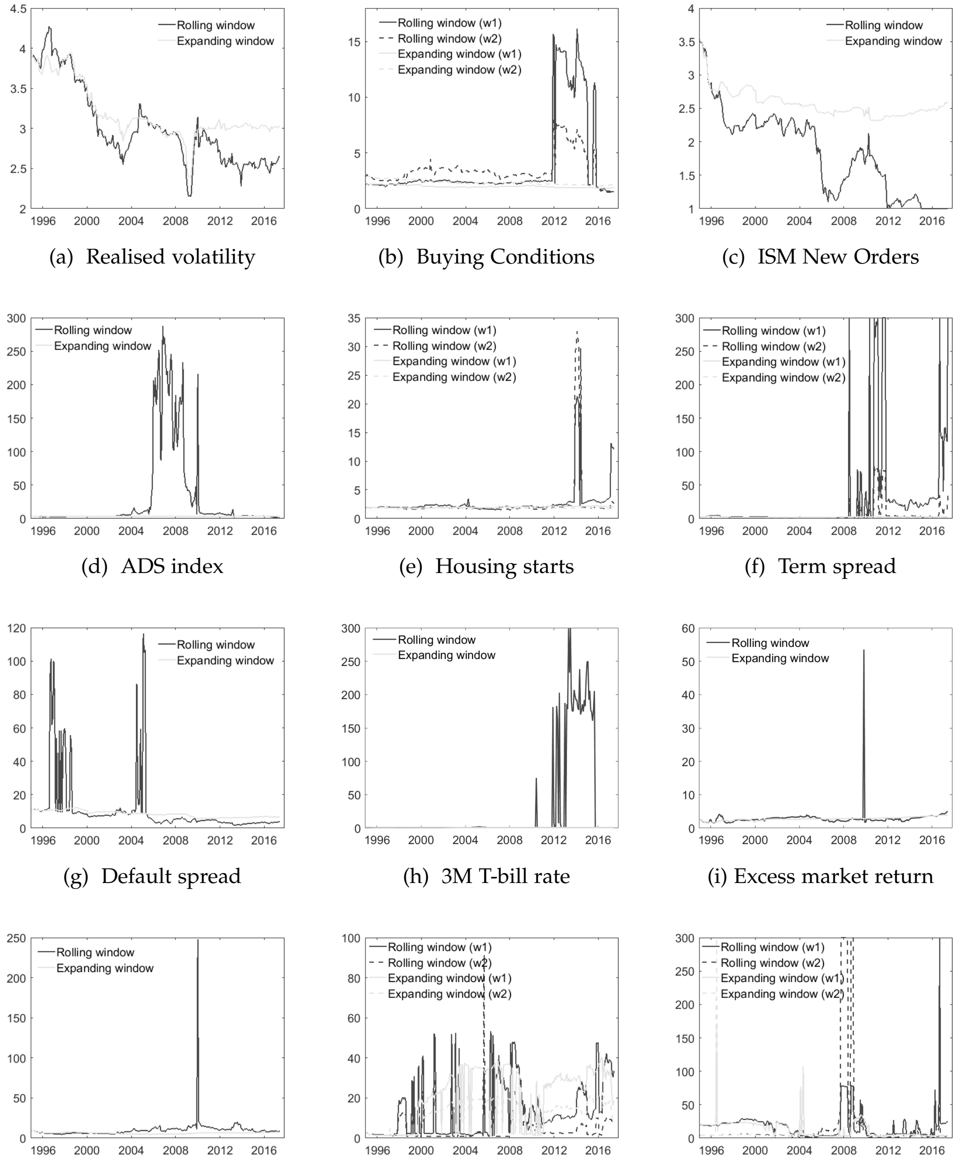
(a) Realised volatility (b) Buying Conditions (c) ISM New Orders

(d) ADS index (e) Housing starts (f) Term spread

(g) Default spread (h) 3M T-bill rate (i) Excess market return

(j) Principal component 1 (k) Principal component 2 (l) Principal component 3

Figure 3.34: t-statistics of the estimated θ parameters from the rolling window GARCH-MIDAS models with fixed weights and the rolling and expanding window GARCH-MIDAS models with the weights re-estimated each period.



(a) Realised volatility (b) Buying Conditions (c) ISM New Orders

(d) ADS index (e) Housing starts (f) Term spread

(g) Default spread (h) 3M T-bill rate (i) Excess market return

(j) Principal component 1 (k) Principal component 2 (l) Principal component 3

Figure 3.35: Estimates of w from the GARCH-MIDAS models estimated using a rolling window and the GARCH-MIDAS models estimated using an expanding window.

To conclude, first note that the statistical tests by Giacomini and White (2006) and Giacomini and Rossi (2010) are not valid for an expanding window estimation scheme, which is why the rolling window scheme is used in the main body of the text. We can, however, conclude that when it comes to out-of-sample forecasting, the rolling window estimation scheme tends to underestimate the usefulness of especially financial data to forecasting stock market volatility, compared to an expanding window estimation scheme. The macroeconomic data and the term spread tend to gain slightly from the flexibility arising from the rolling window estimation scheme. The concern that the volatility arising from re-estimating the weights in every period would significantly impact, and especially worsen, the forecasting performance seems largely unfounded, with only the excess returns and the third principal component gaining clearly from using fixed weights.

3.G Mincer-Zarnowitz regressions

The Mincer-Zarnowitz regressions (Mincer and Zarnowitz, 1969) take the following form: $RV_t = c + \rho\widehat{RV}_{t,t-h} + \epsilon_t$, where $\widehat{RV}_{t,t-h}$ is the forecast from a GARCH model or a GARCH-MIDAS model. If the forecast is unbiased $c = 0$ and $\rho = 1$. In Table 3.15 I test the individual hypothesis that $c = 0$ or $\rho = 1$, as well as use an F-test to test whether c and ρ are jointly equal to 0 and 1, respectively. Clearly, for the 1 and 3 month horizons we cannot reject the null hypothesis of $c = 0$ and $\rho = 1$, and thus the forecasts seem unbiased. On the 6 month horizon the forecasts produced by the GARCH-MIDAS models driven by macroeconomic data, the term spread and the principal components seem unbiased. On the 12 month horizon only the GARCH-MIDAS models driven by housing starts, the term spread and the second principal component produce unbiased forecasts. Looking at the (unadjusted) R^2 statistics performance between the models diverges clearly at the longer horizons, and especially the term spread and housing starts driven GARCH-MIDAS models perform well.

Table 3.15: Mincer-Zarnowitz regressions

	1M ahead				3M ahead				6M ahead				12M ahead			
	c	ρ	F stat	R^2	c	ρ	F stat	R^2	c	ρ	F stat	R^2	c	ρ	F stat	R^2
GJR	-2.85	1.30	456***	64.04	6.07	1.17	25.2***	8.95	15.04	0.82	2.04	0.79	53.73***	-1.28***	2.19	0.85
	(5.30)	(0.25)			(4.42)	(0.24)			(12.13)	(0.75)			(12.70)	(0.56)		
RV	4.97	0.84	273***	51.59	20.99***	0.33***	17.8***	6.49	27.32***	0.11***	1.93	0.75	30.26***	0.01***	0.01	0.00
	(4.10)	(0.20)			(4.90)	(0.09)			(6.75)	(0.06)			(7.21)	(0.05)		
BC	-3.83	1.36	509***	66.52	-4.97	1.53*	68.1***	21.01	-1.43	1.63	29.4***	10.29	22.00***	0.44	1.76	0.68
	(5.21)	(0.25)			(8.04)	(0.50)			(13.15)	(0.82)			(7.90)	(0.40)		
ISM	-2.45	1.32	477***	65.06	3.85	1.35	29.8***	10.42	12.51	0.99	4.76**	1.83	24.70***	0.33*	0.42	0.16
	(5.50)	(0.27)			(4.86)	(0.33)			(10.91)	(0.76)			(8.96)	(0.39)		
HS	-3.83	1.35	496***	65.94	1.56	1.37	42.9***	14.34	-2.69	1.68	27.3***	9.64	-5.21	1.86	28.5***	10.03
	(4.88)	(0.24)			(5.43)	(0.36)			(16.03)	(0.98)			(22.13)	(1.36)		
ADS	-2.85	1.33	378***	59.59	-0.11	1.47	44.0***	14.66	5.43	1.30	12.1***	4.52	22.62***	0.42	0.97	0.38
	(5.76)	(0.29)			(6.97)	(0.45)			(10.78)	(0.72)			(7.37)	(0.44)		
TS	-10.42*	1.79***	499***	66.11	-22.83	2.89*	57.9***	18.45	-33.50	3.64	56.9***	18.19	-31.51	3.58	56.6***	18.10
	(5.60)	(0.29)			(15.29)	(0.98)			(26.17)	(1.67)			(23.64)	(1.57)		
DS	7.68	0.82	294***	53.47	20.10***	0.44***	19.0***	6.92	27.68***	0.12***	1.30	0.51	31.10***	-0.03***	0.06	0.02
	(4.91)	(0.23)			(4.42)	(0.18)			(6.16)	(0.10)			(6.90)	(0.05)		
3MT	-2.13	1.38	468***	64.62	8.48*	1.22	23.2***	8.32	21.20*	0.59	0.93	0.36	36.30***	-0.39	0.24	0.09
	(5.14)	(0.26)			(4.37)	(0.28)			(12.57)	(0.99)			(12.66)	(1.07)		
EMR	-3.74	1.33	248***	49.25	9.71*	0.93	18.1***	6.61	31.78**	-0.06*	0.02	0.01	48.61***	-0.96***	5.42**	2.07
	(7.64)	(0.39)			(5.72)	(0.30)			(15.70)	(0.61)			(15.15)	(0.56)		
PC1	-1.36	1.22	528***	67.33	7.63*	1.04	33.2***	11.48	16.48**	0.68	6.80***	2.59	24.42***	0.30**	1.14	0.44
	(4.78)	(0.22)			(4.57)	(0.27)			(8.22)	(0.49)			(7.03)	(0.31)		
PC2	-2.70	1.42*	493***	65.81	0.85	1.66*	42.3***	14.17	-14.52	2.83	30.2***	10.57	0.65	1.96	9.17***	3.46
	(4.84)	(0.25)			(5.11)	(0.37)			(23.23)	(1.66)			(15.40)	(1.25)		
PC3	5.05	1.05	285***	52.70	11.71***	0.98	29.5***	10.35	14.52*	0.92	7.17***	2.72	24.19***	0.38	0.74	0.29
	(6.01)	(0.31)			(3.76)	(0.20)			(7.65)	(0.55)			(8.78)	(0.61)		

*Parameter estimates for MZ regressions estimated using OLS. HAC standard errors can be found in parenthesis below the parameter estimates. *, ** and *** indicate a rejection of the null hypothesis (i.e., that $c = 0$ or $\rho = 1$, or for the F statistic that jointly $c = 0$ and $\rho = 1$) at the 10%, 5%, and 1% level, respectively. The used variables with their abbreviations are: GJR-GARCH(1,1) (GJR), Realised volatility (RV), Buying Conditions index (BC), ISM New Orders index (ISM), ADS index (ADS), Housing starts (HS), Term spread (TS), Default spread (DS), 3M T-bill rate (3MTb), Excess market return (EMR), Principal component (PC).*

3.H Additional results on the effect of the economic environment

Figures 3.36 and 3.37 illustrate how the NBER recession dates, industrial production growth, VIX index and the St. Louis Fed Financial Stress Index is divided into recession versus expansion and high versus low volatility periods.

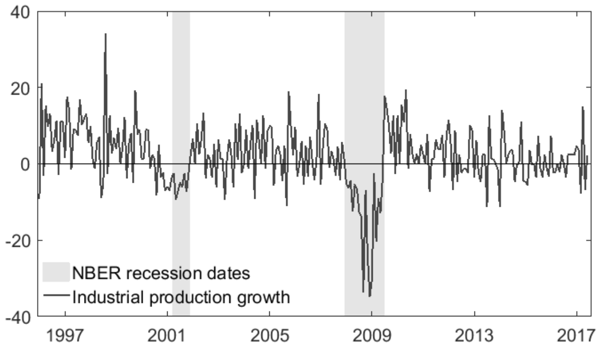


Figure 3.36: NBER recession dates and industrial production growth. Zero is the cut-off point for industrial production growth.

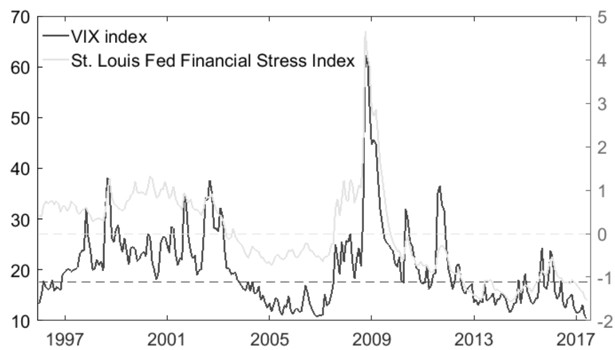


Figure 3.37: VIX index (LHS) and St. Louis Fed Financial Stress Index (RHS). Dashed lines denote the cut-off point of high versus low volatility (or financial stress) periods for each series.

There has only been two recessions during the sample period (from March 2001 to November 2001 (8 months) and from December 2007 to June 2009 (18 months)) and two longer episodes of negative industrial production growth, but several shorter spells of negative growth. The VIX index divides the out-of-sample period into roughly four episodes when using the median as the cut-off point: high volatility from 1996 to 2003, low volatility from 2003 until the beginning of the financial crisis in 2007, the financial crisis and its aftermath, and the largely low volatility since roughly 2013. The St. Louis

Fed Financial Stress Index, which is defined so that zero is the cut-off point between high and low stress regimes, divides the data roughly similarly.

Table 3.16: Effect of business cycle (NBER) on forecasting performance

	1 month ahead				3 months ahead				6 months ahead				12 months ahead			
	Expansion		Recession		Expansion		Recession		Expansion		Recession		Expansion		Recession	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
BC	1.00	1.00	1.01	0.96	0.99	1.01	0.91**	0.88	0.99	1.00	0.90*	0.89	0.99	1.02	0.96	0.96*
ISM	0.99	1.00	1.00	0.99	0.99	1.03	0.99	0.99	0.99	1.03	0.99*	0.99	0.95	0.97	1.01	1.00
HS	1.00	1.00	0.96	0.96	1.02	1.03	0.95	0.94	1.01	1.03	0.91**	0.89	0.98	0.99	0.88**	0.87*
ADS	0.98	0.98	1.13	1.21	1.01	1.01	0.99	0.94	1.02	1.02	0.94**	0.94*	1.01	1.00	0.98	0.98
TS	1.01	1.05	1.09	1.33	0.96	1.02	1.04	1.00	0.90***	0.97	0.93**	0.93	0.85***	0.91*	0.90***	0.91*
DS	1.02	1.01	1.26	1.36	1.13	1.36	1.16	1.04	1.28	1.97	1.05	0.98	1.32	2.13	0.99	0.99
3MT	1.02	1.03	1.00	1.05	1.01	1.08**	1.03***	1.02**	1.01	1.09**	1.03***	1.02**	1.00	1.08*	1.02***	1.01***
EMR	0.99	1.01	1.32**	1.59	1.01	1.01	1.03	1.01	1.04	0.98	1.04***	1.03	1.09**	1.08	1.04***	1.03*
RV	1.02	1.02	1.42*	1.39	1.13	1.37	1.40*	1.25	1.30	2.19	1.26	1.19	1.48	3.19	1.01	0.97
PC 1	0.99	0.99	0.99	0.81	1.02	1.00	0.98	0.95	1.06	1.09	0.95**	0.94	1.05	1.04	0.97	0.97
PC 2	1.01	1.03	1.00	1.04	0.98	1.06	0.98	0.97	0.95	1.06	0.97	0.95	0.94*	1.01	0.97	0.98
PC 3	1.02	1.04	1.12	1.33	0.98	1.04	1.02	0.99	0.96	1.03	1.00	0.98	0.99	1.06	0.98**	0.99*

*Benchmark: GJR-GARCH(1,1) mode. MSFE ratio: $\frac{MSFE_{GMX}}{MSFE_{GARCH}}$, where $MSFE_{GMX}$ stands for the mean squared forecast error of the GARCH-MIDAS-X model. The MAFE ratio is calculated equivalently. A value below 1 means the GARCH-MIDAS-X model outperforms the GJR-GARCH(1,1) model, and vice versa. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test. Recession months are defined according to the NBER Business Cycle Dating Committee. Abbreviations of variables: BC is the Buying conditions index, ISM is the ISM New Orders index, HS denotes housing starts, ADS the ADS index, TS is the term spread, DS the default spread, 3MT the 3 month T-bill rate, EMR is the excess market return, realised volatility: $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$, and PC denotes principal component.*

Table 3.17: Effect of financial market stress on forecasting performance

	1 month ahead				3 months ahead				6 months ahead				12 months ahead			
	Low		High		Low		High		Low		High		Low		High	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
BC	0.98	0.97	1.02	0.98	0.94**	0.98	0.98	0.90	0.95	1.02	0.96	0.90	1.02	1.14	0.96	0.96*
ISM	0.95**	0.97	1.01	1.00	0.92**	0.97	1.04**	1.00	0.92	1.01	1.02*	1.00	0.92	0.95	1.00	1.00
HS	0.97	0.98	1.00	0.97	0.97	0.97	1.01	0.96	0.95	1.00	0.99	0.92	0.94	0.93	0.94	0.90
ADS	0.96	0.96	1.06	1.16	0.96	0.96	1.02	0.96	0.98	1.01	1.00	0.96	0.99	0.90	1.00	0.98
TS	0.97	1.01	1.07	1.27	0.92	1.02	1.02	1.00	0.86**	1.02	0.94**	0.93	0.79***	0.94	0.90***	0.91
DS	0.96*	0.97	1.16	1.29	0.99	0.98	1.22	1.14	1.21	1.83	1.19	1.14	1.55	3.40	1.04*	1.01
3MT	0.98	1.00	1.03*	1.06*	0.94**	1.00	1.06***	1.04***	0.93*	1.04	1.06***	1.04***	0.92	1.04	1.05***	1.03***
EMR	0.95	0.99	1.16**	1.45*	1.00	0.99	1.02	1.02	1.06	0.95	1.03**	1.03	1.16**	1.15	1.03	1.02
RV	0.96	0.97	1.25**	1.32	1.04	1.07	1.31*	1.31	1.31	1.95	1.27	1.37	2.01	6.17	0.97	0.95**
PC 1	0.96	0.98	1.00	0.86	0.95	0.89*	1.03	0.97	1.02	1.01	1.03	0.97	1.07	1.08	0.99	0.97
PC 2	0.97	1.00	1.03	1.05	0.92*	1.05	1.01	0.99	0.88**	1.08	1.00	0.97	0.85**	0.98	1.00	0.99
PC 3	1.02	1.03	1.08	1.26	0.95	1.03	1.02	0.99	0.90**	1.05	1.01	0.99	0.97	1.13	1.00	0.99

*Benchmark: GJR-GARCH(1,1) model. MSFE ratio: $\frac{MSFE_{GMX}}{MSFE_{GARCH}}$, where $MSFE_{GMX}$ stands for the mean squared forecast error of the GARCH-MIDAS-X model. The MAFE ratio is calculated equivalently. A value below 1 means the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model, and vice versa. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test. High / low financial stress months are defined according to the St. Louis Fed Financial Stress Index: 115 high stress months and 143 low stress months. For abbreviations of variable names see notes on Table 3.16.*

3.I Impact of economic environment on parameter estimates

The effect of an economic variable could depend on whether we are in a recession or expansion period, or in a low or high financial stress environment. To incorporate the effect of the economic environment directly in the GARCH-MIDAS model I modify the MIDAS polynomial in equation (3.2) with a dummy variable:

$$log\ \tau_t = m + (\theta_X + \theta_D D_{t-1}) \sum_{k=1}^{K} \varphi_k(\omega_1, \omega_2) X_{t-k} \tag{3.6}$$

where $D_{t-1} = 1$ during a recession, when industrial production growth is negative or when financial conditions[40] are tighter than average.

Table 3.18: GARCH-MIDAS-X-D model: estimates of θ

	IP dummy			NBER dummy			ANFCI dummy		
	θ_X	θ_D	VR	θ_X	θ_D	VR	θ_X	θ_D	VR
Buying Conditions	-0.1742***	-0.0119	14.27	-0.1513***	-0.1409**	17.28	-0.1651***	-0.0373	14.75
ISM New Orders	-0.0506***	0.0020	15.55	-0.0477***	0.0039	16.02	-0.0471***	0.0060***	17.15
Housing starts	-0.0138***	-0.0027	17.39	-0.0120***	-0.0093**	19.64	-0.0148***	-0.0005	17.70
ADS index	-0.3778***	-0.2249*	15.38	-0.4332***	-0.1372	15.68	-0.4846***	0.0066	15.01
Term spread	-0.2632***	0.0664	14.62	-0.2507***	0.3607	15.50	-0.2702**	0.1209*	16.43
Default spread	0.4643***	0.1433**	13.59	0.4714***	0.1479*	14.19	0.3537**	0.1796*	13.72
3M T-bill rate	0.0303**	0.0239*	4.55	0.0291**	0.0450***	6.49	0.0127	0.0427***	7.19
Excess market return	0.1034***	0.0360	10.52	0.1254***	-0.0587	8.91	0.0870***	0.0505	10.43
Realised volatility	0.0622***	0.0078*	36.93	0.0629***	0.0079	38.16	0.0571***	0.0127***	38.77
First PC	0.7652***	0.3688	16.63	0.8939***	0.1264	16.54	0.9964***	-0.1048	15.88
Second PC	-2.1041*	0.4535	10.37	-1.5912***	-1.3232	11.31	-1.3059**	-1.1060	10.26
Third PC	1.1547***	-0.1536	10.73	0.8434***	1.0874**	11.62	0.7643**	0.8026*	10.68

, ** and * indicate statistical significance at the 10%, 5%, and 1% level, respectively. ANFCI index: positive values indicate that financial conditions are tighter than average, and, therefore, $\theta_D = 1$ when the index is positive. NBER: $\theta_D = 1$ when there is a recession. Industrial production growth: $\theta_D = 1$ when annualised monthly growth is negative. $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$. VR is the variance ratio from Section 3.3.1 multiplied by 100. Full results for the models are available upon request.*

The parameter estimates tend to be larger (either positive or negative), and therefore the effect stronger, when the economy is weak or financial conditions are tight. Thus economic data tends to be more important for volatility during weak economic phases or tight financial conditions. None of the dummies is statistically significant for all the variables. The dummy based on industrial production growth is rarely strongly significant and mostly important for financial data. The recession dummy is important for both some macroeconomic and financial variables, while the financial conditions dummy is mostly significant for the financial data.

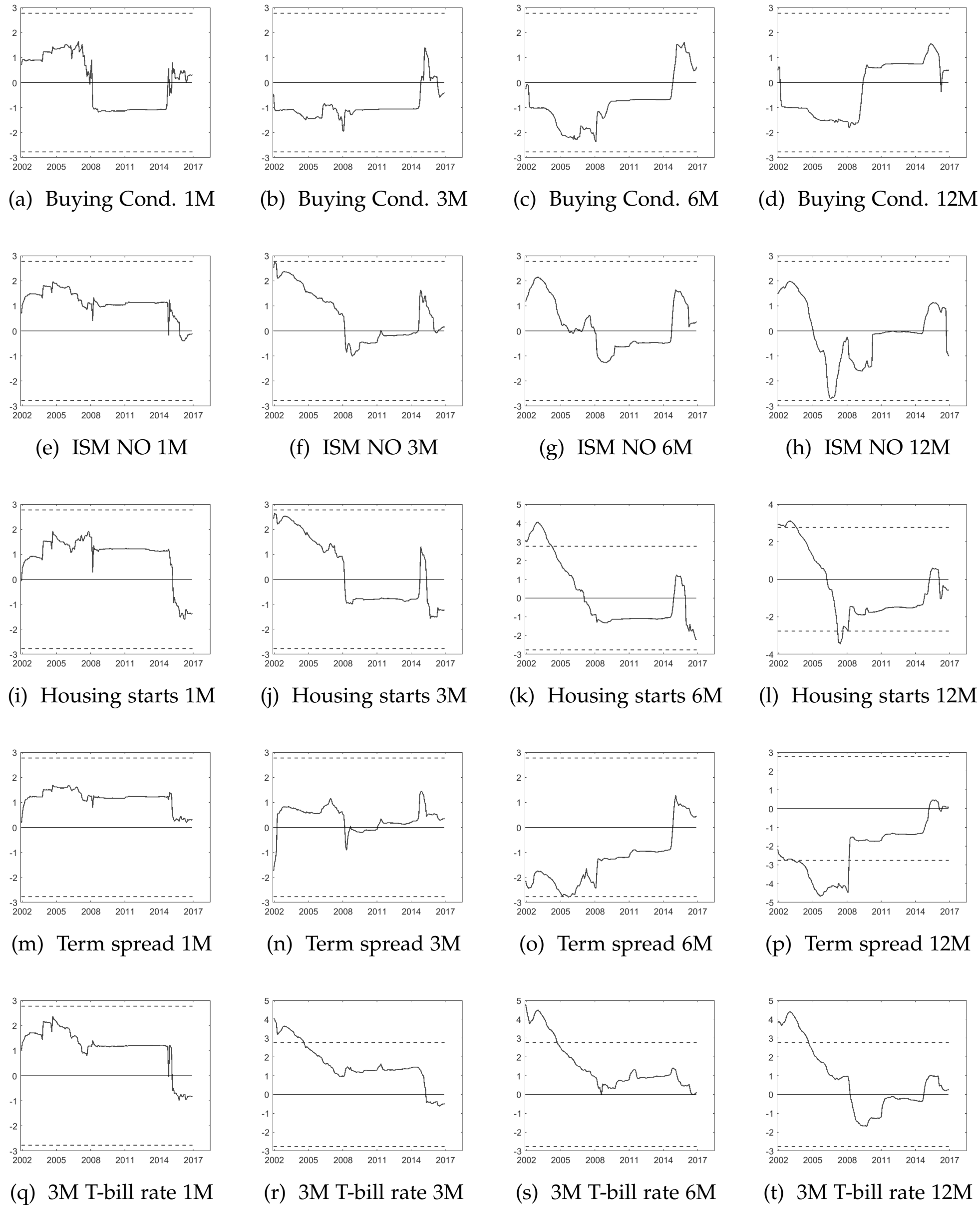
[40] Financial conditions are measured by the ANFCI, which is the Chicago Fed Adjusted National Financial Conditions index. This index was chosen instead of, for example, the VIX index because it has a long enough history to cover the whole sample period.

As the NBER dummy is only available at a considerable lag, I use the financial conditions dummy for the out-of-sample forecasts. As can be seen from Table 3.19, forecasting performance is not significantly improved from Table 3.3, and is even worse in some cases. The Fluctuation test results in Figure 3.38 confirm this conclusion: allowing the effect of the macroeconomic data to be different in strong and weak financial environments improves forecasting performance for, for example, the Buying Conditions index 3M ahead and the 3M T-bill rate, but worsens performance for, for example, the ISM New Orders index 1M, 3M and 6M ahead and 1M ahead for housing starts. The evidence in favour of allowing the impact of the economic data to depend on financial conditions when forecasting volatility seems, therefore, relatively weak.

Table 3.19: Forecasting performance of the GARCH-MIDAS-X-D model

	1 month ahead		3 months ahead		6 months ahead		12 months ahead	
	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE	MAFE	MSFE
Buying Conditions	1.01	0.90	0.98	0.85	1.01	0.92	1.05	1.04
ISM New Orders	1.02	1.09	1.01	1.01	0.99	0.98	0.98	1.00
Housing starts	1.03*	1.07	1.00	0.98	0.99	0.96	0.96*	0.97
ADS index	1.04	1.17	1.00	0.97	1.00	0.97	1.02	1.00
Term spread	1.04	1.23	0.98	1.01	0.91***	0.94	0.88***	0.93
Default spread	1.09	1.22	1.12	1.10	1.19	1.21	1.20	1.29
3M T-bill rate	1.01	1.07	1.02	1.03**	1.01	1.02**	1.00	1.01
Excess market return	1.10	1.59	1.04	1.04	1.04*	1.04	1.08**	1.05
Realised volatility	1.20	1.50*	1.25	1.50	1.30	1.66	1.34	1.72
First PC	0.99	0.88	1.00	0.96	1.02	0.97	1.02	0.98
Second PC	1.01	1.03	0.98	1.00	0.97	0.98	0.96*	0.99
Third PC	1.03*	1.03*	1.01	1.02**	1.01	1.01	1.03	1.03

Benchmark: GJR-GARCH(1,1) model. MSFE ratio: $\frac{MSFE_{GMX}}{MSFE_{GARCH}}$*, where* $MSFE_{GMX}$ *stands for the mean squared forecast error of the GARCH-MIDAS-X-D model (ANFCI dummy). MAFE ratios calculated equivalently. A value below 1 means the GARCH-MIDAS model outperforms the GJR-GARCH(1,1) model. *, ** and *** indicate a rejection of the null hypothesis of equal (unconditional) predictive ability at the 10%, 5%, and 1% level, respectively, according to the Giacomini and White (2006) test.* $RV_t = \sum_{i=1}^{N_t} |r_{i,t}|$.



(a) Buying Cond. 1M (b) Buying Cond. 3M (c) Buying Cond. 6M (d) Buying Cond. 12M

(e) ISM NO 1M (f) ISM NO 3M (g) ISM NO 6M (h) ISM NO 12M

(i) Housing starts 1M (j) Housing starts 3M (k) Housing starts 6M (l) Housing starts 12M

(m) Term spread 1M (n) Term spread 3M (o) Term spread 6M (p) Term spread 12M

(q) 3M T-bill rate 1M (r) 3M T-bill rate 3M (s) 3M T-bill rate 6M (t) 3M T-bill rate 12M

Figure 3.38: Fluctuation test results for loss function differences between the GARCH-MIDAS-X-D model and the GJR-GARCH(1,1) model. Squared forecast errors. Dashed lines represent 90% confidence bands. The year on the x-axis marks the end of the rolling window period over which the test statistic is calculated.

4 Economic origins of the dynamic co-movement of exchange rate and equity returns

4.1 Introduction

The co-movement of financial markets is important for investors but is also of interest to policy makers. Understanding the causes of changing correlations is thus important. Although the causal relationship between the equity market and the foreign exchange (FX) market has been widely studied (see Section 4.2), the economic foundation of the dynamic co-movement between these markets has received less attention. However, as can be seen from Figure 4.1, the correlation varies clearly over time, which motivates studying the dynamic relationship between the two markets.

There are two main theories for the long-term co-movement of equities and exchange rates. First, exchange rate movements can be related to the equity market through the current account, i.e., exchange rate variations affect the competitiveness of international companies and hence their stock price (for example, Dornbusch and Fischer (1980)). The second explanation, which I focus on in this paper, relates to the international portfolio rebalancing channel, i.e., that the value of the stock market affects the demand for money and interest rates, which in turn influences capital flows and therefore the currency (for example, Hau and Rey (2005)).

Clearly equity returns and FX returns are interlinked; the question is rather whether the correlation is positive or negative. While in the current account based view the sign of the correlation depends on the relative importance of exporters and importers, in the portfolio rebalancing framework there are arguments both in favour of a positive correlation (traditional view, which states that a stronger equity market implies stronger currency) and a negative correlation (see Section 4.2 and Hau and Rey (2005)). Cho et al. (2016) concluded that the sign of the correlation seems country-specific: positive for emerging markets and negative for developed countries. From Figure 4.1 it is clear that the sign of the correlation depends on the time period considered, and that correlations calculated over the whole sample period (dashed lines) give a limited picture of the nature of the correlations. The clear time-variation suggests it is worthwhile to consider the economic drivers of the correlation.

Hau and Rey (2005) argued that the correlation is driven by net equity flows, which

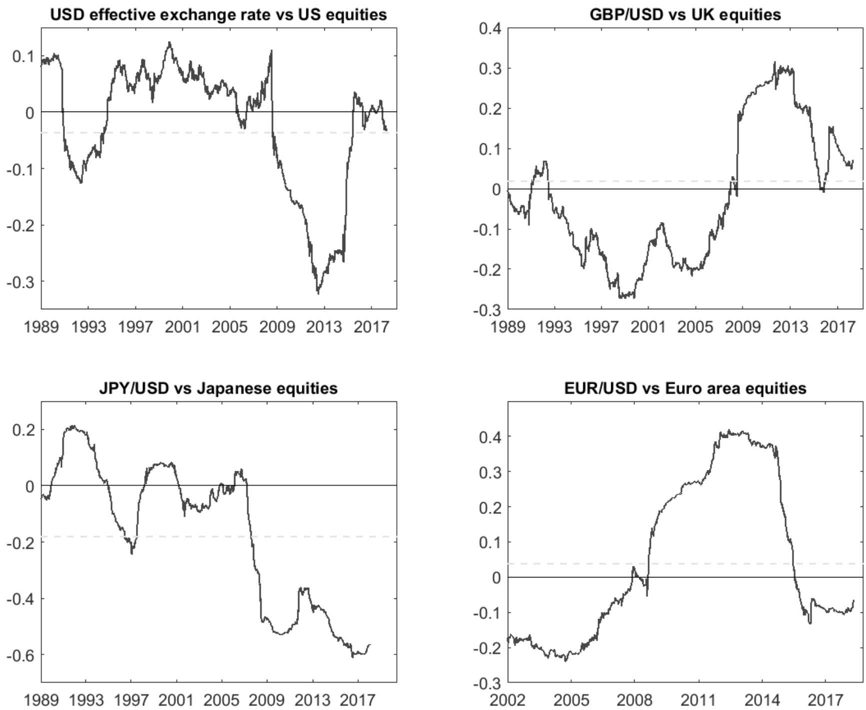


Figure 4.1: Rolling window correlation over 3 years (792 days) between exchange rate and equity returns. The dashed line is the unconditional correlation.

arise from differences in the relative performance of equity markets. In this paper I investigate the ability of macro-finance variables to explain the time-varying correlation, arguing that international investors rebalance their investment portfolios in response to macroeconomic or financial market developments, which lead to net equity flows affecting the exchange rate. A similar idea was explored in Moore and Wang (2014) and Kryzanowski et al. (2017), who used a dynamic conditional correlation (DCC) model to extract the correlation between equity and FX returns. In a second step they used a linear regression model to relate the extracted time-varying correlation to interest rate differentials and trade balance data (Moore and Wang, 2014) or the Federal Reserve's three quantitative easing (QE) programmes (Kryzanowski et al., 2017). Both found a significant relationship between the correlation and their explanatory variables, but the effect varied across markets and QE programmes.

In this paper I use the DCC-MIDAS model[1], which is a two-component model for dynamic correlations, to study the time-variation in the foreign exchange - equity mar-

[1]Dynamic Conditional Correlation model (Engle, 2002) combined with a MIxed DAta Sampling framework (Ghysels et al., 2004), which allows combining data measured at different frequencies in a parsimonious way within the same model. The DCC-MIDAS model was first introduced by Colacito et al. (2011). See Section 4.3 for details.

ket correlation. The model allows combining daily equity and FX return data with monthly macro-finance data within the same model, thus avoiding a two-step estimation procedure and an aggregation of the financial data to a common frequency with the macroeconomic data. As economic data is used to extract the low-frequency trend of the correlation, economic fundamentals can be linked directly to this slow-moving component. This paper thus provides novel evidence on why and how the equity and FX markets co-move by considering several explanatory variables linked to the portfolio rebalancing channel. I examine whether the sign of the correlation depends on the economic environment, and consider both the flight-to-quality channel suggested by Cho et al. (2016) and the search-for-yield channel arising from quantitative easing by central banks, studied earlier by Kryzanowski et al. (2017). I discuss the results for four major developed markets (the US, the UK, Japan and the Euro area) in order to gauge whether the same variables drive the correlations in all regions.

My foremost general conclusion is that central bank unconventional monetary policy and long-term interest rate differentials are the most important drivers of the long-term correlation, although the effect, and therefore the theoretical foundation, varies across markets. Both the flight-to-quality and the search-for-yield channel is supported in the US, as there is strong evidence that a weaker economic or market environment as well as quantitative easing leads to an increasingly negative correlation, while the correlation can be positive during strong growth phases and in low volatility environments. In the US particularly the first and second QE programmes have impacted the correlation negatively. In Japan, where the Yen is an important funding currency, the correlation is driven by interest rate differentials: when the interest rate differential (to the US) shrinks and the motives to use the Yen as a funding currency get weaker the correlation becomes increasingly negative.

In the UK and the Euro area measures for quantitative easing are the most important drivers of the correlation, but the sign is reversed to what we would expect according to the portfolio rebalancing hypothesis. As the equity markets in the UK and the Euro area are clearly correlated with the US equity market, the results might be influenced by developments in the US. In the Euro area the correlation is primarily positive over the sample period, in line with the traditional portfolio rebalancing view of a strong equity market going together with a strong currency, and vice versa. Overall the results for quantitative easing suggest actual purchases are more important than announcement effects for the long-term correlation. An interesting finding is that the choice between actual returns and returns in excess of the US market does not matter for the conclusions from a qualitative point of view, and that for Japan and the Euro area neither does the choice between the spot rate (against the USD) and the effective exchange rate.

This paper is organised as follows. In Section 4.2 I discuss the relevant literature concerning the portfolio rebalancing channel and the relationship between the equity market and the currency market. Section 4.3 presents the DCC-MIDAS model and Section 4.4 the data to be used. Section 4.5 discusses the estimation results, while Section 4.6 concludes the paper.

4.2 Relationship between equity and FX markets

The portfolio balance approach, where changes in the stock market affect the currency market through the current account, was already discussed in, for example, Branson (1983) and Frankel (1983). If domestic stock prices decline the wealth of domestic investors falls, which lowers the demand for money and hence suppresses the interest rate, which in turn encourages capital outflows and hence currency depreciation. This traditional view therefore implies a positive correlation between domestic equities and the domestic currency, regardless of the state of the macroeconomy or the financial markets. The model presented in Moore and Wang (2014), based on Malliaropulos (1998), also predicts a positive relationship between equities and FX.[2] However, also in a portfolio rebalancing framework, Hau and Rey (2005) argued that incomplete risk sharing implies the correlation is negative. They developed an equilibrium model in which equity markets and currency markets are linked through capital flows. Differences in the relative performance of home and foreign equity markets lead to changes in the exchange rate risk exposure of investors, which gives rise to international portfolio rebalancing, and thereby exchange rate movements. Hau and Rey (2004) confirmed, using a VAR with exchange rates, stock markets and net equity flows, that equity returns and portfolio rebalancing are important sources of exchange rate fluctuations.

An alternative perspective is offered by Cho et al. (2016), who studied the correlation between equity markets and exchange rates in both developed and emerging markets. The correlation arises from capital flows, which are induced by for example flight-to-quality behaviour. They concluded that the correlation is positive for emerging markets and negative for developed markets, with equity market conditions influencing net capital flows only in down markets.

In addition, the large scale asset purchases, or quantitative easing (QE) programmes, undertaken by several central banks during the latest financial crisis, can affect the pattern of capital flows through the portfolio rebalancing channel (see, for example, Bernanke (2012)). If different financial assets are not considered perfect substitutes, then the asset purchases by a central bank, which expand its balance sheet, can be expected to raise the price and lower the yield of other financial assets, such as equities. If the portfolio rebalancing undertaken by investors as a response to quantitative easing leads to increased equity flows between countries, then the portfolio rebalancing channel would lead to an exchange rate depreciation. QE would therefore induce a negative correlation between the equity market and the foreign exchange market.

In addition to these channels, Bae and Elkamhi (2016) suggested that if the financial markets are sufficiently integrated, a common risk factor, reflecting the same aggregate risk, could drive the risk premia on both markets. Pavlova and Rigobon (2007) found, utilising an international asset pricing model including demand shocks and trade in goods, that the exchange rate is important for determining both equity and bond market dynamics.

The empirical literature on the long-term relationship between equity markets and exchange rates mostly rely on cointegration analysis and error correction models to determine a causal relationship between the equity market and the currency market.

[2]Note that in Moore and Wang (2014) this is a negative correlation as the exchange rate is expressed as USD/'home'.

This is of interest because the current account based hypothesis indicates that causality runs from exchange rates to equities, while the portfolio rebalancing channel indicates the opposite is true. Granger causality tests are commonly used to differentiate between the two hypotheses. The literature reaches mixed conclusions, at times finding no stable long-term relationship between the two (for example, Bahmani-Oskooee and Sohrabian (1992) for the US, Nieh and Lee (2002) for G7 countries), while other studies document a significant (causal) relationship between stock markets and exchange rates (in either or both directions) (for example, Kim (2003) for the US, Ajayi and Mougoué (1996) for eight advanced economies, and, for example, Phylaktis and Ravazzolo (2005) and Lin and Fu (2016) for Asian markets). Kollias et al. (2012) found time-varying causality with data from the Euro area. Lin (2012) (Asia), Caporale et al. (2014) (developed markets) and Tsagkanos and Siriopoulos (2013) (the EU and the US) compared the linkages of stock markets and FX markets in crisis and non-crisis periods, finding varying causal relationships. Cenedese et al. (2015) concluded, based on a cross-sectional portfolio strategy, that there is no relationship between stock returns and exchange rates on the aggregate level, and recommend that research should concentrate on individual countries and specific shocks.

While the DCC-MIDAS model has not been used before to study the correlation between equity and FX returns, Kryzanowski et al. (2017) used a DCC model to extract the correlation betwen bond markets, stock markets and currency forwards for a broad range of countries. In a linear regression framework with time dummies for quantitative easing (QE) periods they studied how correlations have changed during the Federal Reserve's three QE programmes. They found that while QE1 (11/2008–3/2010) had on average no effect on correlations, QE2 (11/2010–6/2011) lead to higher and QE3 (9/2012–10/2014) to lower correlations compared to no-QE periods, and that the effect was similar in both developed and emerging markets. Whether QE has impacted the relationship between different (international) financial markets has been studied by, for example, Belke et al. (2017) and Thornton (2014).

Moore and Wang (2014) also used a DCC model to extract the dynamic correlation between FX and equities, and subsequently a linear regression framework to explain the correlation using trade balance data and interest rate differentials (on mortgage rates). They found that trade balance data is important for Asian markets, suggesting the current account based explanations are important, while the interest rate differential is important for developed markets, such as the UK, supporting the portfolio balance hypothesis.

4.3 The DCC-MIDAS model

Engle (2009) outlined three steps needed to specify and estimate a dynamic conditional correlation (DCC) model. The first step consists of estimating the volatilities, which are needed to construct the standardised (i.e., volatility-adjusted) residuals. In the second step, based on the standardised residuals, the dynamic quasi-correlations are estimated. In the last step the quasi-correlation matrix is rescaled to ensure that it is a proper correlation matrix. The same steps are required to estimate the DCC-MIDAS model, introduced by Colacito et al. (2011). The DCC model can be estimated by quasi maximum likelihood (QML) methods. The resulting QML estimator will be consistent albeit inefficient if the mean and covariance models are correctly specified (Engle, 2009). To my knowledge no estimation theory has been developed for the DCC-MIDAS model, but the model is estimated using QMLE in, for example, Conrad et al. (2014) and Asgharian et al. (2016).

The DCC-MIDAS model combines the DCC model of Engle (2002) with the GARCH-MIDAS model of Engle et al. (2013). It decomposes both the volatilities and the correlation into a short-term component and a long-term trend. The short-term component of the correlation follows a DCC scheme and the long-term trend reflects past realised correlations. The DCC-MIDAS-X extension, introduced by Conrad et al. (2014), allows the long-term component to depend on economic and financial data. The slowly moving long-term component, around which the short-term component fluctuates, can therefore be thought of as reflecting the fundamental economic causes of time variation in the correlation. The mixed data sampling framework allows for the direct use of monthly economic data in a daily model for correlations.

For estimating the first step volatilities I use the GARCH-MIDAS model, where the returns for each asset on day t and in month τ can be modelled as having a multiplicative specification for the conditional variance:

$$r_{t,\tau} = \mu + \sqrt{m_{t,\tau}\, g_{t,\tau}}\, \varepsilon_{t,\tau}, \quad \varepsilon_{t,\tau} \mid \Phi_{t-1,\tau} \sim N(0,1), \quad \forall t = 1, ..., N_\tau \tag{4.1}$$

where $\Phi_{t-1,\tau}$ represents the information set up to day $t-1$, and N_τ is the number of trading days in period τ. $\sigma^2_{t,\tau} = m_\tau\, g_{t,\tau}$ is the total conditional variance, where m_τ[3] represents the (monthly) long-term volatility component and $g_{t,\tau}$ the (daily) GARCH component. The daily $g_{t,\tau}$ component can, for example, follow an asymmetric GJR-GARCH(1,1) model (Glosten et al. (1993)):

$$g_{t,\tau} = 1 - \alpha - \beta - \gamma/2 + (\alpha + \gamma D_{t-1,\tau}) \frac{(r_{t-1,\tau} - \mu)^2}{m_\tau} + \beta g_{t-1,\tau} \tag{4.2}$$

where $\alpha + \beta + \gamma/2 < 1$, $\alpha > 0$, $\beta \geq 0$ and $\alpha + \gamma \geq 0$. $D_{t-1,\tau}$ is an indicator function, taking the value 1 when $(r_{t-1,\tau} - \mu) < 0$ and 0 otherwise. Thus the parameter γ describes the degree of asymmetry in volatility. When $\gamma = 0$ the standard GARCH(1,1) model is obtained.

The long-term component is described by a MIDAS polynomial:

$$m_\tau = \bar{m}_v + \theta_v \sum_{k=1}^{K_v} \varphi_k(\omega_v) RV_{\tau-k} \tag{4.3}$$

where K_v is the number of lags of realised volatility ($RV_\tau = \sum_{t=1}^{N_\tau} r^2_{t,\tau}$) included and

[3] $m_{t,\tau}$ is fixed for all t in period τ. Hence, the subscript t is suppressed to ease notation and emphasise that m_τ evolves at the lower frequency than $g_{t,\tau}$.

$\varphi_k(\omega_v)$ is a weighting scheme following a beta polynomial[4]:

$$\varphi_k(\omega_v) = \frac{(1 - \frac{k}{K_v})^{\omega_v - 1}}{\sum_{j=1}^{K_v}(1 - \frac{j}{K_v})^{\omega_v - 1}}, \quad \text{where} \sum_{k=1}^{K_v} \varphi_k(\omega_v) = 1. \tag{4.4}$$

For the second step correlations we can first note that the return vector follows the process $r_t \sim N(\mu, \mathbf{H_t})$. Following Engle (2002) the conditional covariance matrix can be decomposed as $\mathbf{H_t} = \mathbf{D_t R_t D_t}$, where $\mathbf{R_t}$ is the conditional correlation matrix of the standardised return residuals and $\mathbf{D_t}$ is a diagonal matrix with the standard deviations of the returns on the diagonal. Then, $\mathbf{R_t} = E_{t-1}[\xi_t \xi_t]$, where $\xi_t = \mathbf{D_t^{-1}}(r_t - \mu)$ are the standardised residuals obtained from the GARCH-MIDAS model. The short-term quasi-correlations are then estimated as:

$$q_{12,t} = \bar{\rho}_{12,\tau}(1 - a - b) + a(\xi_{1,t-1}\xi_{2,t-1}) + bq_{12,t-1} \tag{4.5}$$

where $\xi_{1,t-1}$ and $\xi_{2,t-1}$ are the standardised residuals from the GARCH-MIDAS models of assets 1 and 2. These correlations fluctuate around a long-term time-varying trend ($\bar{\rho}_{12,\tau}$). According to Engle (2009) $q_{12,t}$ can be thought of as an approximation of the true conditional correlation. However, as the diagonal elements do not necessarily equal exactly one a rescaling of the conditional correlation matrix is necessary: $\mathbf{R_t} = diag\{\mathbf{Q_t}\}^{-1/2}\mathbf{Q_t}diag\{\mathbf{Q_t}\}^{-1/2}$.

In the DCC model by Engle (2002) $\bar{\rho}_{12,\tau} = \bar{\rho}_{12}$, i.e., the intercept is time-invariant. The DCC-MIDAS model by Colacito et al. (2011) assumed that $\bar{\rho}_{12,\tau}$ is time-varying and that it can be expressed as a weighted sum of past realised correlations (RC):

$$\bar{\rho}_{12,\tau} = \sum_{k=1}^{K_c} \varphi_k(\omega_c) RC_{\tau-k}, \quad RC_\tau = \frac{\sum_{t=N_{\tau-1}+1}^{N_\tau} \xi_{1,t}\xi_{2,t}}{\sqrt{\sum_{t=N_{\tau-1}+1}^{N_\tau} \xi_{1,t}^2 \sum_{t=N_{\tau-1}+1}^{N_\tau} \xi_{2,t}^2}}. \tag{4.6}$$

The DCC-MIDAS-X model by Conrad et al. (2014) allows $\bar{\rho}_{12,\tau}$ to depend on economic data. To ensure the conditional covariance remains positive (despite potentially negative economic data) the Fisher-z transformation is used:

$$\bar{\rho}_{12,\tau} = \frac{exp(2\bar{z}_{12,\tau}) - 1}{exp(2\bar{z}_{12,\tau}) + 1}, \tag{4.7}$$

where

$$\bar{z}_{12,\tau} = \bar{m}_c + \theta_c \sum_{k=1}^{K_c} \varphi_k(\omega_c) X_{\tau-k}. \tag{4.8}$$

Here K_c is the number of lags of the explanatory data (X_τ) included in the long-term component and $\varphi_k(\omega_c)$ is a weighting scheme, as in equation (4.4).

The DCC-MIDAS model can be estimated by a two-step procedure, as described

[4]The beta polynomial could also be estimated with two free parameters: ω_1 and ω_2. However, it has been established in the earlier literature (see, for example, Engle et al. (2013)) that realised volatility tends to have a decaying weighting scheme, i.e., $\omega_1 = 1$. The restricted weighting scheme is chosen in the interest of parsimony.

in Engle (2002) and Colacito et al. (2011). Following Engle (2002) the log-likelihood function can be written as:

$$LLF = -\frac{1}{2}\sum_{t=1}^{T}\left(n\ log(2\pi) + 2\ log|\mathbf{D_t}| + r_t'\mathbf{D_t^{-2}}r_t\right) - \frac{1}{2}\sum_{t=1}^{T}\left(log|\mathbf{R_t}| + \xi_t\mathbf{R_t^{-1}}\xi_t + \xi_t'\xi_t\right) \tag{4.9}$$

From this expression it is clear why the model can be estimated in two steps: the first part of the likelihood function only depends on the data and parameters of the GARCH-MIDAS model, while the second part of the likelihood function depends on the standardised residuals and the DCC parameters. The first part of the likelihood function can, therefore, be used to estimate the variance parameters, and in a second step, conditional on these GARCH-MIDAS model parameter estimates, the parameters relating to the correlation model can be estimated from the second part of the likelihood function.[5]

4.4 Data and hypotheses

My data set includes the United States (US), the United Kingdom (UK), Japan and the Euro area. The sample period is from January 1986 to June 2018, except for the Euro area, for which the sample starts in January 1999. All data sources are reported in Appendix 4.A. The daily equity market data is the S&P500 index for the US, the Nikkei 225 index for Japan, the FTSE index for the UK and the STOXX50 for the Euro area. The equity market data is in domestic currency. I use both actual returns and returns in excess of the US market return (hereafter, excess returns). Hau and Rey (2005) and Moore and Wang (2014) argued in favour of using excess returns while Cho et al. (2016) used actual returns. While excess returns have been cleaned of common global variation, actual returns can be important when studying, for example, flight-to-quality behaviour. For the daily exchange rate I use the effective exchange rate (EER), and for countries other than the US I also use the bilateral spot exchange rate against the USD. For the spot rate I use the convention 'home'/USD, i.e., a positive correlation between home currency returns and home equity market returns arises when the home currency appreciates (depreciates) and the home equity market rises (falls). Log returns are used for both exchange rates and equities. Table 4.1 presents descriptive statistics of the return data. Equity returns tend to be on average positive while excess returns are negative, i.e., the US return on average exceeds the return on the other equity markets. The mean returns on the FX markets are close to zero, and they vary less than equity returns.

I examine three different hypotheses and therefore three different sets of variables for explaining the correlation between equities and FX returns. First of all, I consider

[5]Note that this two-step estimation affects the standard errors of the parameters. While the Bollerslev-Wooldridge robust standard errors can be used when estimating the GARCH-MIDAS model parameters, the formula for the standard errors of the DCC-MIDAS model is much more complicated, see Engle and Sheppard (2001) for the formula. However, following the earlier empirical literature Bollerslev-Wooldridge robust standard errors are used here for the second step as well.

Table 4.1: Descriptive statistics for the daily return data

Foreign exchange returns	Mean	Standard deviation	Minimum	Maximum	N
USD EER	-0.0035	0.4068	-3.7517	2.8073	8066
GBP EER	-0.0042	0.4343	-6.1001	2.6345	8068
JPY EER	0.0071	0.6408	-5.9998	6.4038	7730
EUR EER	-0.0001	0.3767	-2.9225	2.8732	4878
GBP/USD	0.0004	0.6448	-8.1597	6.1646	8068
JPY/USD	0.0075	0.7018	-3.9871	7.2408	7730
EUR/USD	0.0006	0.6440	-4.7354	4.2207	4878
Equity returns					
US	0.0306	1.1228	-22.8997	10.9572	8066
UK	0.0247	1.0948	-13.0286	9.3843	8068
Japan	0.0097	1.4554	-16.1354	13.2346	7730
Euro area	0.0152	1.3208	-13.0286	10.4376	4878
Excess equity returns					
UK	-0.0135	1.1328	-18.2240	11.4223	7161
Japan	-0.0250	1.7286	-21.3308	20.5195	7161
Euro area	-0.0200	1.2485	-10.6347	14.6379	4307

Sample period: January 1986 - June 2018 or January 1999 - June 2018 (Euro area). N denotes the number of observations.

the flight-to-quality argument, presented by Cho et al. (2016), which predicts a negative correlation between equities and FX in developed markets (or safe havens) and a positive correlation in emerging markets. In particular, Cho et al. (2016) argued that in global down markets capital tends to move out of emerging markets into developed countries, leading to a currency appreciation in the developed country, and therefore a negative correlation between FX and equity returns. I measure flight-to-quality motives by stock market volatility (VXO index) and for the US also by the news-based Economic Policy Uncertainty index by Baker et al. (2016).[6] For example, Bauwens and Otranto (2016) suggested that market volatility is an important driver of correlations.[7]

Second, I consider the search-for-yield channel related to quantitative easing (QE), studied also in Kryzanowski et al. (2017). This relates to the portfolio rebalancing channel of unconventional monetary policy, where the low yields resulting from QE translate into higher prices for other assets as investors shift to higher-yielding assets, including foreign assets. Ex-ante one would expect quantitative easing to lead to a negative correlation between FX and equity returns as investors shift to domestic equities as well as foreign assets. As I consider low-frequency movements in the correlation, I proxy QE by central bank balance sheet data (assets) but also consider time dummies following Kryzanowski et al. (2017) and announcement effect dummies in line with Rogers et al. (2014). As an alternative measure the monetary base (M0) is used.

[6]For the UK the corresponding index starts clearly later than my sample period and in Japan the index has been discontinued. Therefore, the data is only used for the US.

[7]It would be interesting to construct volatility indices for each market, but at the moment the US VXO index is used as a proxy for global market conditions.

Third, I investigate whether the correlation is sensitive to standard macroeconomic variables, such as economic growth or inflation (both actual and relative performance compared to the US). In line with the hypothesis of Hau and Rey (2005), who emphasised relative equity market performance as a driver of the correlation, if (relative) economic growth is indicative of (relative) equity market performance it can also influence the correlation. Cho et al. (2016) concluded that capital flows are sensitive to stock market performance only in down markets, and similarly the correlation might be more sensitive to economic conditions in, for example, recessions. Interest rate differentials (to the US) based on the policy rate and the 10-year yield are also included in order to determine whether the portfolio balance approach is supported.[8]

The explanatory variables are considered on a monthly frequency. Some of them are also available at a higher frequency, but as the focus of this paper is on the variation in the long-term trend component of the correlation, rather than day-to-day fluctuations, all data is aggregated (average over the month) to monthly frequency. The sample period is chosen such that all the data, except the central bank balance sheet data, is available for the entire time period. On the other hand, central bank balance sheets have only attracted attention since the financial crisis, and therefore cannot be expected to provide any information on the correlations before this. Thus the balance sheet data is only considered from 2003 onwards.

4.5 Results

The first subsection reports the unconditional, full-sample correlations between FX and equity returns, while the second section considers the dynamic correlations produced by the DCC-MIDAS-RC model, i.e., the DCC-MIDAS-X model driven by realised correlations.[9] The third section presents the GARCH-MIDAS model estimation results for volatility, while the fourth section discusses the results for the DCC-MIDAS-X model for correlations. The fifth subsection considers quantitative easing dummies, while the last subsection presents a robustness check using a multivariate regression model.

4.5.1 Full-sample correlations

The full sample correlations, presented in Table 4.2, are all small, with the largest correlations being for Japan, where equity returns and FX returns are negatively correlated. For the remaining markets the correlations are slightly positive. On the other hand, the correlations for excess equity returns are all negative. Thus the results confirm the negative correlation between excess equity returns and FX returns presented in Hau and Rey (2005), while the findings in Cho et al. (2016) regarding actual returns are confirmed for Japan.

[8]To incorporate the effects of unconventional monetary policy on the interest rate a shadow rate is also used, but this does not bring any additional insight compared to using the policy rate and the 10-year yield. Results are available upon request.

[9]The estimations are executed in Matlab building on the basic code in the MIDAS Matlab Toolbox.

Table 4.2: Full sample correlation between FX and equity returns

	Effective exchange rate (EER)		Spot rate ('home'/USD)	
	Equity return	Excess equity return	Equity return	Excess equity return
United States	0.0362	-	-	-
United Kingdom	0.0363	-0.0178	0.0175	-0.0408
Japan	-0.2289	-0.1334	-0.1806	-0.1099
Euro area	0.0227	-0.0232	0.0378	-0.0188

Sample period: January 1986 - June 2018 or January 1999 - June 2018 (Euro area).

4.5.2 Dynamic correlations

Figure 4.2 illustrates the dynamic conditional correlation and the long-term correlation component (with 36 lags for the correlations[10]) extracted using the DCC-MIDAS-X model with realised correlation driving the long-term component. The parameter estimates for each model can be found in the tables in Section 4.5.4. Clearly the correlations vary over time and the results are similar to those obtained using the simple rolling window correlation (see Figure 4.1). Appendix 4.A confirms that most of the correlations can be considered unambiguously dynamic.

The remainder of this paper concentrates on the correlation between the spot rate and the actual equity return, except for the US, for which the effective exchange rate is used. This is because, first of all, the results for excess returns are very similar to those using actual returns. Secondly, for Japan and the Euro area the correlations behave similarly regardless of which exchange rate measure is chosen. For the UK the correlation based on the effective exchange rate fluctuates largely around zero, while that of the spot rate varies more clearly. Results for the UK, Japan and the Euro area relating to the EER and to excess equity returns can be found in Appendix 4.E and 4.D.

4.5.3 GARCH-MIDAS model estimation results

This section presents results for the GARCH-MIDAS-RV model for equity and FX returns, which is the first step in the estimation of the DCC-MIDAS-X model. Two years of lagged data (i.e., $K_v = 24$) is used in all models. This adequately captures the lag structure: all weight parameters (ω_v) are such that the weights decline towards zero over the 24 lags (see Table 4.3).[11]

Foreign exchange volatility does not display asymmetric effects for any market (γ is insignificant), while equity returns display, as expected, strong asymmetric effects, with $\gamma > 0$ implying that lower-than-expected returns lead to higher volatility. I thus proceed using the GARCH(1,1) model in the GARCH-MIDAS specification for FX re-

[10]A lag length of three years is considered suitable as focus in this paper is on the medium run. This is a frequency which abstracts from short term noise, but is still short enough to capture some business cycle fluctuations.

[11]The log-likelihood function is maximised at this lag for almost all models. At $K = 12$, which maximises the log-likelihood for the US FX return, weights in the MIDAS polynomial do not decay to zero. For Euro area equity returns the increase in the log-likelihood value at $K = 36$ compared to $K = 24$ is very small. Therefore, in the interest of having an equal number of lags in all models, $K = 24$ is chosen.

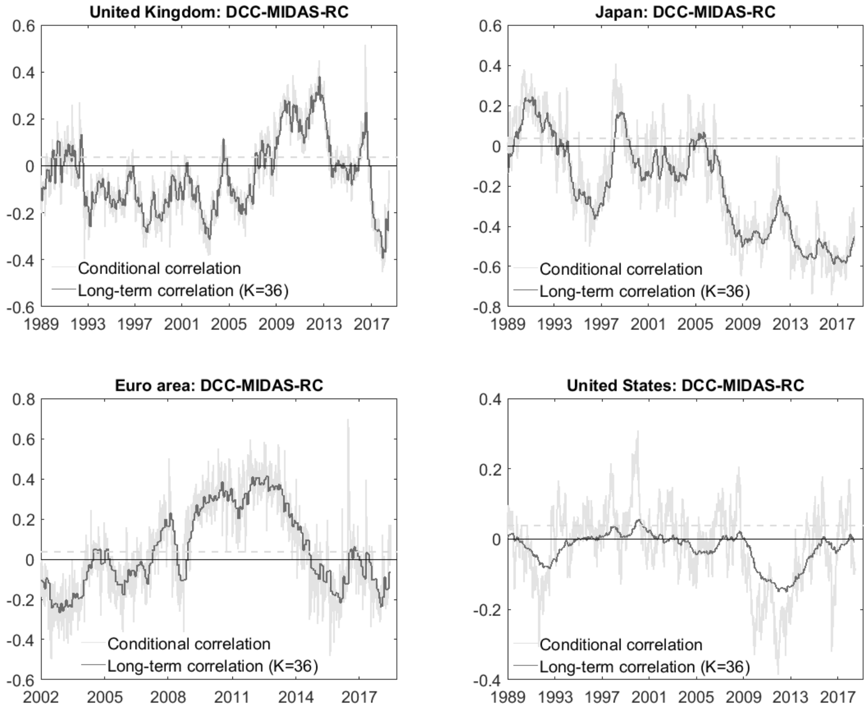


Figure 4.2: Long-term and conditional correlation between equity and FX returns, estimated using the DCC-MIDAS-X model driven by realised correlations.

turns and the GJR-GARCH(1,1) model for equity returns. The estimation results for the chosen models are reported in Table 4.3. Lagged realised volatility gets a highly significant and positive parameter estimate (θ_v) in most models, indicating it is an important driver of long-term volatility.[12] The variance ratios are also relatively high. For example, for US equity returns the long-term volatility component explains more than 40 % of the fluctuations in total volatility. For Euro area FX returns a standard GARCH(1,1) model is used to construct the standardised residuals.

4.5.4 DCC-MIDAS-X model estimation results

The following subsections discuss the estimation results for the DCC-MIDAS-X model, using 36 lags of the macro-finance data outlined in Section 4.4. The models are estimated based on the standardised residuals from the models presented in Table 4.3.

[12]Notice that when testing the significance of θ_v, θ_v and the weight parameters ω_v are not separately identified under the null hypothesis, which affects the asymptotic distribution of the test statistic. However, I follow the convention in the GARCH-MIDAS literature (for example, Engle et al. (2013) and Conrad and Loch (2014)) and proceed using the standard t-statistic. See Ghysels et al. (2007) for a discussion of the problem in MIDAS regressions.

Table 4.3: Estimation results of the GARCH-MIDAS-RV model

	μ	α	β	γ	θ_v	ω_v	$\bar{m}_v$	VR
USD EER returns	-0.0014	0.0553***	0.9102***	-	0.1711***	5.9863**	0.2493***	33.62
	(0.0040)	(0.0097)	(0.0309)		(0.0277)	(2.9002)	(0.0609)	
GBP/USD returns	0.0070	0.0785***	0.8549***	-	0.1846***	5.4825***	0.3157***	56.78
	(0.0058)	(0.0128)	(0.0248)		(0.0133)	(1.3951)	(0.0468)	
JPY/USD returns	-0.0032	0.0699***	0.8807***	-	0.1572***	7.1767	0.4999***	26.21
	(0.0073)	(0.0189)	(0.0740)		(0.0424)	(6.0614)	(0.1397)	
EUR/USD returns	0.0055	0.0272***	0.9694***	-	-	-	0.4188***	-
	(0.0082)	(0.0038)	(0.0043)				(0.0883)	
US equity returns	0.0320***	0.0010	0.7941***	0.2255***	0.1740***	7.4037***	0.5689***	44.25
	(0.0088)	(0.0104)	(0.0266)	(0.0355)	(0.0113)	(1.3574)	(0.0566)	
UK equity returns	0.0269***	0.0247***	0.8126***	0.1688***	0.1619***	8.2306***	0.6346***	37.16
	(0.0092)	(0.0093)	(0.0277)	(0.0342)	(0.0128)	(1.9550)	(0.0529)	
Japan equity returns	0.0349***	0.0449***	0.7801***	0.2113***	0.1914***	7.8952**	0.8212***	31.31
	(0.0127)	(0.0124)	(0.0476)	(0.0580)	(0.0198)	(3.9637)	(0.1148)	
Euro area equity returns	0.0014	0.0010	0.8519***	0.1988***	0.1804***	5.6005***	0.7100***	36.64
	(0.0149)	(0.0153)	(0.0193)	(0.0288)	(0.0136)	(1.6268)	(0.1084)	

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. The MIDAS polynomial: $m_\tau = \bar{m}_v + \theta_v \sum_{k=1}^{K_v} \varphi_k(\omega_v) RV_{\tau-k}$, where RV stands for realised volatility (sum of squared returns). VR stands for the variance ratio, a measure of in-sample fit: $100\frac{Var(log(m_\tau))}{Var(log(m_\tau g_{t,\tau}))}$. The models are estimated for the sample period January 1986 - June 2018, with two years of lagged data used for the first period. For the Euro area the sample period starts in January 1999.*

The most important parameter in the tables in this section is θ_c, which gives the effect of the economic data on the long-term correlation. As Conrad et al. (2014) explained, due to the non-linear nature of the model, the sign of θ_c is directly interpretable, but the marginal effect of the explanatory data on the correlation is not. Note that for a number of the models in this section the weight parameter of the MIDAS polynomial (ω_c) is at either the lower (≈ 1) or the upper (50) bound. A large weight parameter indicates that only the most recent explanatory data matters for the correlation. A weight parameter close to one implies a moving average of the previous 36 lags is included, meaning the weights do not decay to zero. Unsurprisingly, these occur mainly for the quantitative easing measures, for which large changes can plausibly impact the market environment for a long time. For example, Kryzanowski et al. (2017) used dummies which were equal to one during the entire duration of QE and found significantly different correlations in QE and non-QE periods. The moving average weighting scheme reflects a very persistent effect of the explanatory variable on the correlation.[13]

4.5.4.1 The United States

The correlation between the US dollar effective exchange rate returns and the US equity market returns is mostly influenced by the general economic situation (represented by industrial production growth), the VXO index, the Economic Policy Uncertainty (EPU) index, and central bank asset data (or the monetary base), as evidenced by the

[13]More lagged data could be included in the MIDAS polynomial in an effort to render the weighting scheme decaying, but this would also shorten the estimation period.

highly significant θ_c parameter estimates in Table 4.4. Figure 4.3 shows the conditional correlations together with the long-term components extracted using the DCC-MIDAS-X models. The figure confirms that the same above-mentioned five variables are able to capture the sharp shifts in the correlation during the financial crisis, while industrial production growth, the EPU index and the VXO index extract well-fitting trends from the correlation for the whole sample period. In addition, it seems that the VXO index leads the developments in the correlation, making it a promising variable for an out-of-sample exercise.[14] The AIC and the BIC indicate that the DCC-MIDAS-X models driven by macro-finance data lead to a better fit than the DCC-MIDAS-RC model.

Table 4.4: DCC-MIDAS-X model: USD EER returns vs. US equity returns

	a	b	θ_c	ω_c	$\bar{m}_c$	LLF	AIC	BIC
Realised correlation	0.0123***	0.9793***	0.4456*	1.6760	-0.0192	-14438.8	3.5814	3.5857
	(0.0031)	(0.0067)	(0.2566)	(1.4132)	(0.0301)			
Inflation (YoY)	0.0117***	0.9832***	0.0121	15.1706	-0.0592	-14439.7	3.5816	3.5860
	(0.0030)	(0.0047)	(0.0290)	(16.9022)	(0.0830)			
Monetary base (M0) (Δ)	0.0100***	0.9791***	-0.1427***	1.1**	0.0925***	-14425.6	3.5781	3.5825
	(0.0032)	(0.0062)	(0.0240)	(0.4597)	(0.0356)			
Central bank assets (Δ)	0.0097***	0.9802***	-0.1016***	1.1*	0.0340	-14425.8	3.5782	3.5825
	(0.0031)	(0.0061)	(0.0173)	(0.6301)	(0.0288)			
IP growth (MoM)	0.0102***	0.9805***	0.3911***	1.1	-0.0873***	-14428.2	3.5788	3.5831
	(0.0032)	(0.0060)	(0.0846)	(0.7791)	(0.0242)			
EPU index	0.0117***	0.9794***	-0.0026***	4.6455*	0.2321**	-14433.9	3.5802	3.5845
	(0.0030)	(0.0057)	(0.0009)	(2.6675)	(0.0973)			
VXO index	0.0108***	0.9836***	-0.0137***	5.4609	0.2437**	-14433.9	3.5802	3.5845
	(0.0031)	(0.0051)	(0.0053)	(4.6096)	(0.1195)			

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. Δ denotes change-over-period. LLF denotes the log-likelihood function, AIC the Akaike information criterion and BIC the Bayesian information criterion. IP denotes industrial production and EPU denotes Economic Policy Uncertainty.*

The negative parameter estimate (θ_c) on the VXO index and the EPU index imply support for the flight-to-quality hypothesis outlined in Section 4.4: the higher market volatility is, the stronger is the negative correlation between equities and FX for a safe haven market like the US. On the other hand, when market volatility (or political uncertainty) is low and there are no flight-to-quality motives, the correlation between the US equity market return and the US dollar return is small or even positive.

The search-for-yield hypothesis relating to quantitative easing is also supported in the US data, as evidenced by the negative estimate for θ_c on the central bank assets (or the monetary base). This implies that, in line with the ex-ante prediction, when the equity market rises the dollar weakens as search-for-yield behaviour dominates, leading to a negative correlation.[15]

Industrial production growth affects the correlation in a highly significant and positive way. Looking at Figure 4.3 we can see that weak economic performance amplifies the negative correlation, while strong economic performance leads to either a weak

[14]Forecasting correlations in an out-of-sample context is beyond the scope of this paper.

[15]The effect of QE is very sharp, and I therefore consider dummies for QE in Section 4.5.5. It is also evident that the correction towards zero occurring around 2012 is influenced by the weighting scheme, i.e., the large increases in QE dropping out of the moving average after three years.

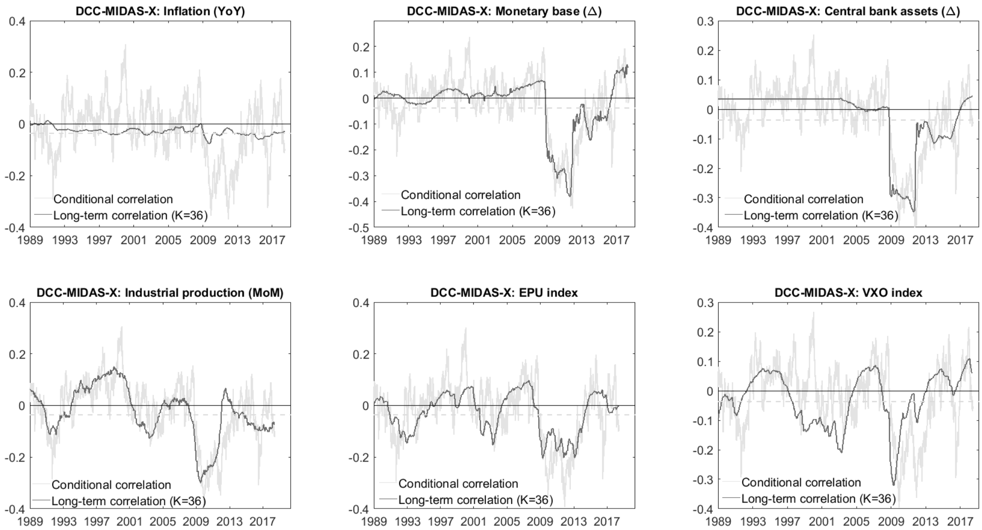


Figure 4.3: Long-term and conditional correlation between the US equity return and the USD EER return.

correlation or a positive correlation. The correlation thus seems sensitive to the economic environment, but naturally this result might also reflect similar drivers as the flight-to-quality or search-for-yield motives.

4.5.4.2 The United Kingdom

We can see from Table 4.5 (specifically parameter θ_c) and Figure 4.4 that similar variables, namely central bank assets (or the monetary base) and the VXO index, are important for the correlation in the UK as in the US. The signs of the estimates for θ_c are, however, reversed. For the UK QE therefore leads to a positive correlation between FX and equity returns, as does higher US market volatility. This is contrary to the logic presented in Section 4.4, unless the UK is interpreted as a market from which capital flees in times of crisis. As the US and UK equity markets are relatively highly correlated (0.48) and the bilateral exchange rate against the USD is used, it is, however, possible that there is an effect from the US feeding through.[16] This interpretation is supported by the fact that when using the GBP effective exchange rate the effect from QE is, if anything, negative (see Appendix 4.E). I also include a US QE measure as an explanatory variable for the UK, in order to consider spillover effects from US monetary policy.[17] US QE gets a weakly significant parameter estimate (θ_c), and looking at Figure 4.4 we can see that it is able to explain the sharp shift in the correlation that

[16]As pointed out by Cho et al. (2016), if equity markets are positively correlated a negative correlation between equity and FX returns in one market may imply a positive correlation in another.

[17]Kryzanowski et al. (2017) documented spillover effects from US unconventional monetary policy to the correlation between equity and FX markets in both developed and emerging countries.

occurred around the financial crisis.

Table 4.5: DCC-MIDAS-X model: GBP/USD return vs. UK equity return

	a	b	θ_c	ω_c	$\bar{m}_c$	LLF	AIC	BIC
Realised correlation	0.0219***	0.9133***	0.8686***	7.2763***	0.0019	-14342.6	3.5567	3.5610
	(0.0050)	(0.0310)	(0.1112)	(1.6180)	(0.0210)			
Policy rate (vs US)	0.0129***	0.9840***	0.0030	5.0153	-0.0475	-14361.8	3.5614	3.5658
	(0.0032)	(0.0045)	(0.0368)	(10.7510)	(0.0758)			
Inflation (YoY)	0.0135***	0.9825***	0.0542	50	-0.1801*	-14358.8	3.5607	3.5650
	(0.0034)	(0.0052)	(0.0410)	(59.4450)	(0.0975)			
Inflation (vs US)	0.0138***	0.9813***	0.0766*	50	-0.0417	-14358.0	3.5605	3.5648
	(0.0037)	(0.0063)	(0.0439)	(51.3360)	(0.0446)			
Long-term yield (vs US)	0.0131***	0.9836***	0.0379	50	-0.0633	-14361.4	3.5613	3.5657
	(0.0033)	(0.0048)	(0.0598)	(83.6680)	(0.0578)			
Monetary base (M0) (Δ)	0.0121***	0.9810***	0.1554***	1.1**	-0.2041***	-14349.1	3.5583	3.5626
	(0.0038)	(0.0065)	(0.0388)	(0.3609)	(0.0448)			
Central bank assets (Δ)	0.0122***	0.9818***	0.1027***	1.1	-0.1257***	-14352.3	3.5591	3.5634
	(0.0036)	(0.0062)	(0.0341)	(0.7244)	(0.0377)			
US central bank assets (Δ)	0.0124***	0.9799***	0.1356*	1.1	-0.1212***	-14349.6	3.5584	3.5627
	(0.0045)	(0.0185)	(0.0697)	(9.8198)	(0.0387)			
IP growth (MoM)	0.0129***	0.9835***	-0.2793	11.1010	-0.0344	-14360.1	3.5610	3.5653
	(0.0032)	(0.0048)	(0.3052)	(11.4100)	(0.0562)			
IP growth (vs. US)	0.0124*	0.9838***	0.4099	3.7339	0.0068	-14359.1	3.5608	3.5651
	(0.0070)	(0.0141)	(0.4395)	(46.2610)	(0.0681)			
VXO index	0.0125***	0.9843***	0.0148**	20.4270	-0.3315**	-14358.6	3.5606	3.5650
	(0.0029)	(0.0041)	(0.0072)	(29.5520)	(0.1390)			

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. Otherwise, see notes on Table 4.4.*

From Figure 4.4 we can see that industrial production growth relative to the US seems able to explain at least some the fluctuations in the correlation in the early part of the sample, with the positive relationship implying that when the UK grows slower than the US the correlation is increasingly negative. Overall, however, the models augmented with macroeconomic data fail to outperform the DCC-MIDAS-RC model in terms of the AIC and the BIC.

4.5.4.3 Japan

From Table 4.6 it is clear that the correlation between the Japanese Yen and the Japanese equity market is clearly driven by the interest rate differential and the inflation differential to the US, as evidenced by the significant parameter estimates on θ_c for these variables. This is not surprising considering the widely acknowledged role the Yen plays as a funding currency in for example carry trade strategies, due to the relatively low interest rate. The estimated coefficients (θ_c) on the inflation and interest rate differentials are negative, implying that, for example, a shrinking interest rate differential to the US leads to an increasingly negative correlation.[18] In practice this means that when the interest rate differential is large (i.e., US rates are clearly above Japanese rates), and therefore the carry trade motives for the Yen are strong, the Japanese equity market return and the Japanese exchange rate return are only weakly correlated.

[18] Note that the interest rate differential between Japan and the US is primarily negative, and a large interest rate differential always means that the interest rate differential is negative.

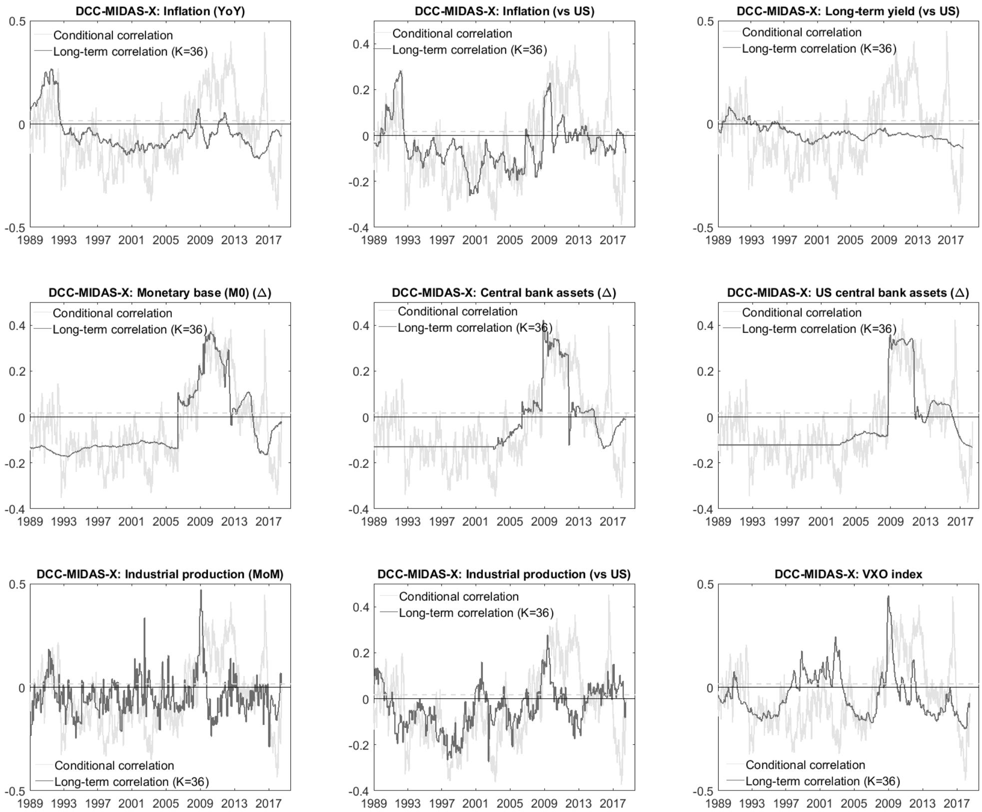


Figure 4.4: Long-term and conditional correlation between the UK equity return and the GBP/USD return.

However, as the interest rate differential to the US shrinks the (negative) co-movement between Japanese equity returns and FX returns strengthens considerably. This implies that when the funding currency motives get weaker (relative to the US) the correlation between currency and equity returns strengthens.

The search-for-yield hypothesis relating to unconventional monetary policy is supported in Japan as the central bank asset data has a highly significant negative impact on the correlation. The positive coefficient (θ_c) on the VXO index does not support a safe haven role for Japan in terms of the flight-to-quality hypothesis, but from Figure 4.5 it is evident that the relationship between the conditional correlation and the VXO index is not very close, for example, during the financial crisis. In terms of the AIC and the BIC the models augmented by macroeconomic data fail to outperform the DCC-MIDAS-RC model.

Table 4.6: DCC-MIDAS-X model: JPY/USD return vs. Japanese equity return

	a	b	θ_c	ω_c	$\bar{m}_c$	LLF	AIC	BIC
Realised correlation	0.0265***	0.9487***	1.0004***	3.5950***	-0.0498	-13143.9	3.4021	3.4066
	(0.0046)	(0.0105)	(0.0759)	(0.8688)	(0.0331)			
Policy rate (vs US)	0.0193	0.9761***	-0.1003	1.1	-0.5060***	-13159.0	3.4060	3.4105
	(0.0179)	(0.0442)	(0.1732)	(20.7580)	(0.2177)			
Inflation (YoY)	0.0178**	0.9797***	0.0371	22.2330	-0.3141***	-13166.5	3.4079	3.4124
	(0.0079)	(0.0098)	(0.0519)	(38.5530)	(0.0873)			
Inflation (vs US)	0.0190***	0.9774***	-0.1586***	1.7881*	-0.5933***	-13161.1	3.4065	3.4110
	(0.0072)	(0.0094)	(0.0422)	(0.9252)	(0.0968)			
Long-term yield (vs US)	0.0192**	0.9758***	-0.2470***	1.1	-0.9150***	-13157.1	3.4055	3.4100
	(0.0080)	(0.0131)	(0.0799)	(4.0803)	(0.1926)			
Monetary base (M0) (Δ)	0.0181**	0.9794***	-0.0194	1.1	-0.2556***	-13166.4	3.4079	3.4124
	(0.0078)	(0.0097)	(0.0205)	(1.2270)	(0.0990)			
Central bank assets (Δ)	0.0199***	0.9754***	-0.1930***	1.1	-0.1727***	-13160.3	3.4063	3.4108
	(0.0068)	(0.0096)	(0.0392)	(1.2351)	(0.0566)			
IP growth (MoM)	0.0177**	0.9799***	0.2828	2.0035***	-0.3063***	-13166.0	3.4078	3.4123
	(0.0086)	(0.0106)	(0.2242)	(0.5241)	(0.0859)			
IP growth (vs. US)	0.0179**	0.9797***	-0.0774	6.7726**	-0.3035***	-13166.6	3.4079	3.4124
	(0.0075)	(0.0093)	(0.1645)	(2.7397)	(0.0746)			
VXO index	0.0186**	0.9786***	0.0237**	50***	-0.7443***	-13161.0	3.4067	3.4112
	(0.0073)	(0.0092)	(0.0093)	(10.6020)	(0.2007)			

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. Otherwise, see notes on Table 4.4.*

4.5.4.4 Euro area

The central bank balance sheet data has a positive and (weakly) statistically significant effect on the long-term correlation in the Euro area (Table 4.7), echoing the results for the UK.[19] While for the Euro area the results extend to using the effective exchange rate, it is noteworthy that also the US and Euro area equity market returns are positively correlated over the sample period (0.56). The US central bank asset data has a significant effect on the Euro area correlation, and when looking at Figure 4.6 it is clear that US QE had a significant, albeit short-lived, impact on the correlation in late 2008, around the time the first round of QE was launched in the US. It is, however, also evident that contrary to the UK, balance sheet expansion in the Euro area and the US had distinct impacts on the correlation. Overall the search-for-yield hypothesis related to QE is thus not directly supported in the Euro area.

Based on Figure 4.6 the long-term interest rate differential seems like an important driver of the correlation as well. The positive parameter estimate (θ_c) indicates that higher (relative) long-term interest rates in the Euro area lead to a positive correlation, which could, for example, imply that higher relative interest rates signal better economic prospects, a higher stock market and therefore a stronger currency in the Euro area, in line with the traditional portfolio balance view. On the other hand, a negative interest rate differential, i.e., higher interest rates in the US, implies a small or negative correlation. The flight-to-quality hypothesis is not supported in the Euro area, as evidenced by the insignificant parameter estimate for the VXO index. Echoing the results

[19] Altavilla et al. (2015) concluded that the Euro depreciated (against the USD) and the Euro area stock market rose on quantitative easing announcement days, implying a negative correlation. However, here I consider the impact over a longer horizon.

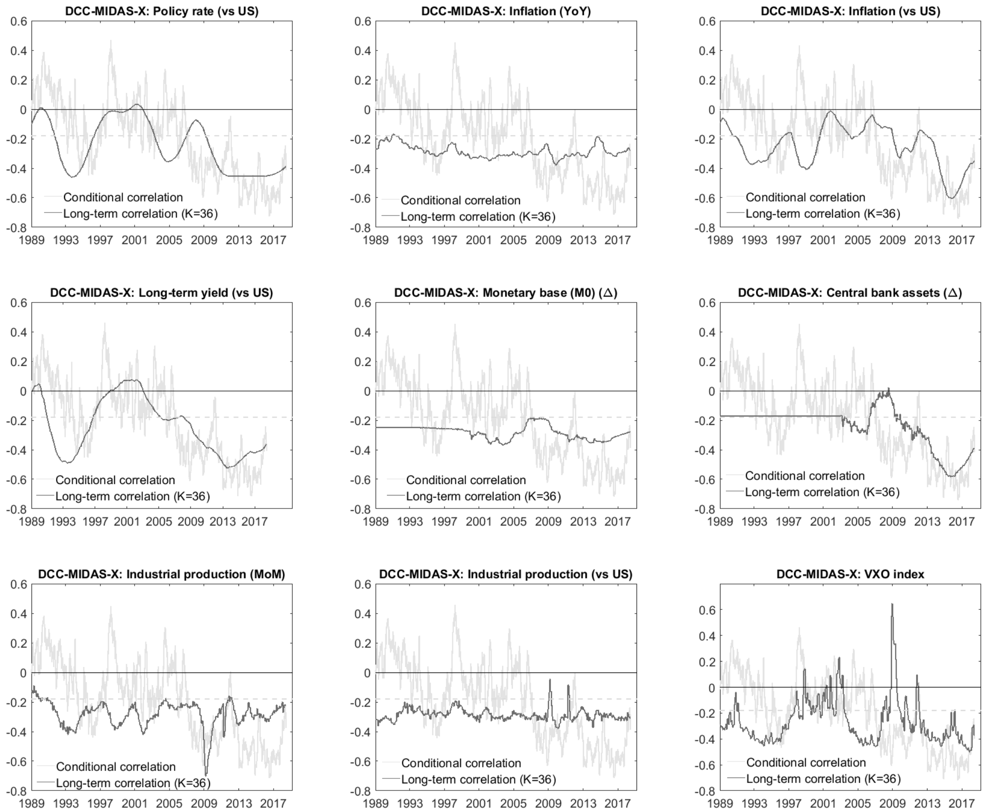


Figure 4.5: Long-term and conditional correlation between the Japanese equity return and the JPY/USD return.

for the UK and Japan the DCC-MIDAS-RC model seems difficult to beat in terms of AIC and BIC.

4.5.5 Quantitative easing dummies

Instead of central bank balance sheet data it is also possible to use dummies for quantitative easing. To make the results comparable to Kryzanowski et al. (2017) I use the same QE dummies for the US, i.e., QE1 between November 2008 and March 2010, QE2 from November 2010 to June 2011, and QE3 from September 2012 to October 2014. In addition, a QE normalisation dummy is included, which covers the time period from June 2017, when the Federal Reserve announced details about its plans for reducing the balance sheet over time, until the end of the sample in May 2018. The balance sheet of the Federal Reserve started contracting gradually in October 2017. The dummies used in Kryzanowski et al. (2017) cover the whole QE period and can therefore be used to differentiate between QE and non-QE periods. Contrary to the balance sheet data this

Table 4.7: DCC-MIDAS-X model: EUR/USD return vs. Euro area equity return

	a	b	θ_c	ω_c	$\bar{m}_c$	LLF	AIC	BIC
Realised correlation	0.0502***	0.8509***	0.8543***	7.3056***	0.0310	-7898.8	3.2406	3.2473
	(0.0156)	(0.0469)	(0.1434)	(2.6387)	(0.0287)			
Policy rate (vs US)	0.0189**	0.9761***	0.0763	7.5463	0.0958	-7916.0	3.2476	3.2543
	(0.0089)	(0.0123)	(0.0743)	(12.8310)	(0.0831)			
Inflation (YoY)	0.0185*	0.9770***	0.0253	4.3988	0.0722	-7917.1	3.2481	3.2548
	(0.0099)	(0.0131)	(0.1594)	(3.8520)	(0.3045)			
Inflation (vs US)	0.0191**	0.9758***	0.1299	5.5684	0.1544***	-7916.3	3.2478	3.2544
	(0.0093)	(0.0130)	(0.1597)	(4.1435)	(0.0963)			
Long-term yield (vs US)	0.0235	0.9645***	0.1680	20.2460	0.0755*	-7911.9	3.2459	3.2526
	(0.0396)	(0.0760)	(0.1049)	(99.1480)	(0.0406)			
Monetary base (M0) (Δ)	0.0175*	0.9782***	0.0849	3.1730	0.0322	-7915.7	3.2475	3.2542
	(0.0103)	(0.0144)	(0.1264)	(3.3200)	(0.1334)			
Central bank assets (Δ)	0.0181*	0.9759***	0.1528*	1.9704	-0.0310	-7913.0	3.2464	3.2531
	(0.0110)	(0.0170)	(0.0925)	(2.3582)	(0.0609)			
US central bank assets (Δ)	0.0184	0.9727***	0.1438**	1.1	-0.0694	-7908.7	3.2446	3.2513
	(0.0127)	(0.0206)	(0.0713)	(0.6803)	(0.0811)			
IP growth (MoM)	0.0176**	0.9780***	-0.3061	9.7089	0.1392*	-7914.0	3.2468	3.2535
	(0.0081)	(0.0109)	(0.4330)	(28.0040)	(0.0715)			
IP growth (vs. US)	0.0172**	0.9792***	0.7973	1.1	0.1009	-7916.11	3.2477	3.2543
	(0.0080)	(0.0105)	(0.9547)	(1.8231)	(0.1063)			
VXO index	0.0173	0.9782***	0.0174	11.0070	-0.2224	-7914.1	3.2469	3.2535
	(0.0161)	(0.0231)	(0.0280)	(60.4020)	(0.5070)			

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. Otherwise, see notes on Table 4.4.*

allows differentiating between the impact of the three quantitative easing programmes.

Alternatively announcement effect dummies for QE can be considered. To this end, announcement dummies identified in Rogers et al. (2014) as being announcements directly related to quantitative easing are included for all four markets.[20] I include four separate dummy variables for the US (QE1, QE2, QE3 and normalisation) in order to determine the impact of each QE programme, while for the other markets one dummy variable covers all the QE announcements. As this is a low-frequency study the announcement effect dummies are included at a monthly level and in the low-frequency component of the DCC-MIDAS model.[21] This can be justified by arguing that from a portfolio rebalancing perspective the effect on the markets might not be instantaneous.

I consider two modifications to the MIDAS polynomial (equation (4.8)) in order to include QE dummies. In the first case I only consider a dummy variable, i.e., the conditional correlation is constant except for the QE dummy:

$$\bar{z}_{12,\tau} = \bar{m}_c + \sum_{i=1}^{4} \theta_{D_i} D_{i,\tau}, \tag{4.10}$$

where $i = 1$ for the UK, Japan and the Euro area. A more interesting case is when the

[20] I.e., LSAP (Large-Scale Asset Purchase) announcements for the US, APF (Asset Purchase Facility) announcements for the UK, APP (Asset Purchase Programme) announcements for Japan, and LTRO (Longer-Term Refinancing Operation) and bond purchase announcements for the Euro area.

[21] Announcement effects could also be studied by including the dummy in the short-term correlation component, but that is outside the scope of this paper.

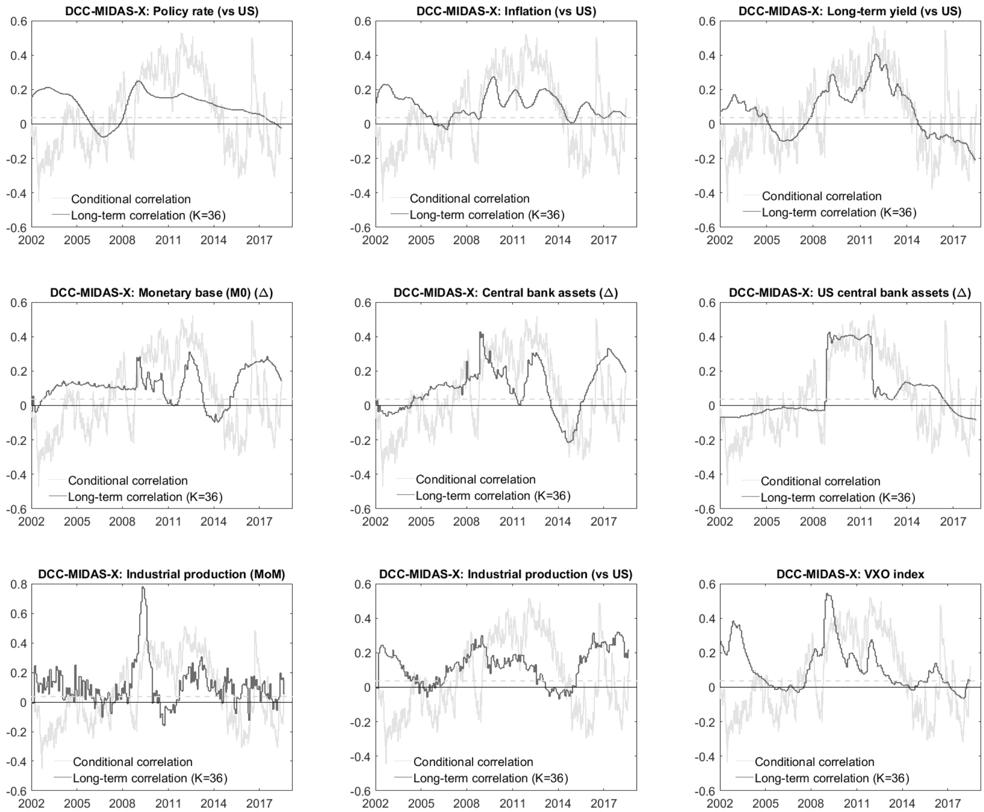


Figure 4.6: Long-term and conditional correlation between the Euro area equity return and the EUR/USD return.

QE dummy is included together with economic data, i.e.,

$$\bar{z}_{12,\tau} = \bar{m}_c + \theta_c \sum_{k=1}^{K_c} \varphi_k(\omega_c) X_{\tau-k} + \sum_{i=1}^{4} \theta_{D_i} D_{i,\tau}. \tag{4.11}$$

Table 4.8 considers the impact of US quantitative easing on the the FX and equity return correlation for the four markets. Only results using the QE dummies of Kryzanowski et al. (2017) are presented, as the announcement dummies by Rogers et al. (2014) are mostly statistically insignificant, implying it is the QE period rather than the actual announcements which affect the long-term correlation.[22] For the US the dummies relating to the first and second QE tend to be statistically significant and negative, while those relating to the third QE programme and the monetary policy normalisation period are not. The QE dummies reduce the significance of the economic data, highlighting the importance of unconventional monetary policy for correlations.

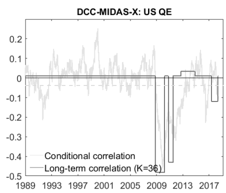
[22]The full results using the QE announcement effect dummies are available upon request.

Table 4.8: DCC-MIDAS-X model: US quantitative easing dummies

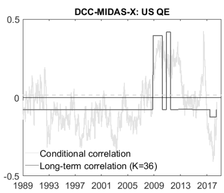
	a	b	θ_C	ω_C	θ_{D_1}	θ_{D_2}	θ_{D_3}	θ_{D_4}	$\bar{m}_C$	LLF	AIC	BIC
United States												
Only QE	0.0079**	0.9861***	-	-	-0.5357***	-0.4706**	0.0240	-0.1301	0.0098	-14426.8	3.5789	3.5850
	(0.0036)	(0.0051)			(0.0874)	(0.1956)	(0.1418)	(0.1944)	(0.0335)			
RC	0.0087***	0.9813***	0.6013	1.1	-0.4314***	-0.2010***	0.1310	-0.0728	-0.0006	-14424.7	3.5789	3.5867
	(0.0032)	(0.0053)	(0.9184)	(3.3480)	(0.0639)	(0.0570)	(0.3275)	(0.1456)	(0.0258)			
IP growth	0.0057	0.9897***	0.2254*	7.6472*	-0.3867***	-0.6151***	0.0779	-0.1397	-0.0352	-14422.9	3.5785	3.5863
	(0.0050)	(0.0067)	(0.1179)	(4.1871)	(0.1493)	(0.2746)	(0.1539)	(0.1913)	(0.0397)			
Inflation	0.0074*	0.9867***	-0.0295	41.0937	-0.6196***	-0.5066**	-0.0181	-0.1585	0.0940	-14425.4	3.5791	3.5869
	(0.0040)	(0.0056)	(0.0284)	(105.3644)	(0.1262)	(0.2124)	(0.1425)	(0.1967)	(0.0948)			
CB assets (Δ)	0.0055	0.9906***	-0.1911**	5.7383***	0.2107	-0.2635	0.3659*	-0.2421	0.0379	-14424.5	3.5788	3.5866
	(0.0041)	(0.0052)	(0.0829)	(1.4453)	(0.3607)	(0.2350)	(0.2115)	(0.2206)	(0.0406)			
M0 (Δ)	0.0072**	0.9871***	-0.1332	3.8370***	-0.0080	-0.2715	0.1577	-0.2059	0.0779	-14424.1	3.5787	3.5866
	(0.0036)	(0.0054)	(0.0829)	(1.3711)	(0.3084)	(0.2011)	(0.1658)	(0.2105)	(0.0610)			
VXO index	0.0075**	0.9871***	-0.0072	17.2308	-0.4518***	-0.5017**	-0.0120	-0.2268	0.1526	-14425.0	3.5790	3.5868
	(0.0036)	(0.0052)	(0.0059)	(35.0065)	(0.1259)	(0.2160)	(0.1481)	(0.2236)	(0.1244)			
EPU index	0.0085***	0.9833***	-0.0231**	1.1*	-0.4388***	-0.2353*	0.1249	-0.1114	0.2231**	-14424.0	3.5787	3.5865
	(0.0018)	(0.0023)	(0.0117)	(0.6372)	(0.0989)	(0.1378)	(0.1565)	(0.1102)	(0.1079)			
United Kingdom												
Only QE	0.0122***	0.9833***	-	-	0.4955	0.5234	0.0026	-0.0480	-0.0773*	-14356.1	3.5605	3.5666
	(0.0039)	(0.0063)			(0.3952)	(0.6155)	(0.2102)	(0.2144)	(0.0446)			
RC	0.0217***	0.9123***	0.8428***	6.8046***	0.0797	0.0924	-0.0640	-0.0208	-0.0011	-14339.7	3.5569	3.5647
	(0.0051)	(0.0361)	(0.1572)	(1.7484)	(0.1192)	(0.1598)	(0.0818)	(0.0746)	(0.0262)			
Japan												
Only QE	0.0176*	0.9797***	-	-	-0.0460	0.5118	-0.3390**	-0.1407	-0.2592**	-13164.6	3.4079	3.4142
	(0.0105)	(0.0136)			(0.2478)	(0.7463)	(0.1679)	(0.4362)	(0.1266)			
RC	0.0255***	0.9479***	1.0264***	3.3084***	0.0880	0.2180***	-0.0936*	0.0817	-0.0486	-13139.6	3.4020	3.4101
	(0.0046)	(0.0109)	(0.0841)	(0.8134)	(0.0557)	(0.0825)	(0.0515)	(0.0677)	(0.0319)			
Euro area												
Only QE	0.0176	0.9763***	-	-	0.4892	0.4014	-0.0359	-0.2415	0.0575	-7911.2	3.2465	3.2558
	(0.0152)	(0.0231)			(0.7389)	(0.7971)	(0.3643)	(0.2402)	(0.0798)			
RC	0.0524***	0.8314***	0.9043***	4.0579*	0.1922	0.0243	-0.1239	-0.1406**	0.0302	-7889.3	3.2383	3.2503
	(0.0158)	(0.0527)	(0.2194)	(2.1372)	(0.1359)	(0.2177)	(0.1332)	(0.0668)	(0.0364)			

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. RC denotes realised correlation. The dummies for QE1, QE2 and QE3 follow the dates identified in Kryzanowski et al. (2017). The QE4 dummy signifies the normalisation period. Otherwise, see notes on Table 4.4.*

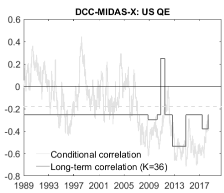
The results for the other markets suggest US quantitative easing has not had a significant effect on the correlation, contrasting the results obtained using the balance sheet data in Section 4.5.4 for the UK and the Euro area. On the other hand, from Figure 4.7 we can see that despite the lack of statistical significance the start of QE1 does seem to coincide with the positive shifts in the correlation for both the UK and the Euro area. The second and third QE programmes seem to have had some impact on the correlation in Japan, while the normalisation period is important in the Euro area. The results for Japan are broadly in line with those in Kryzanowski et al. (2017), where the average correlation in developed markets was higher during QE2 and lower during QE3 compared to non-QE periods.



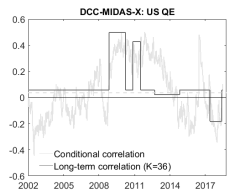
(a) United States



(b) United Kingdom

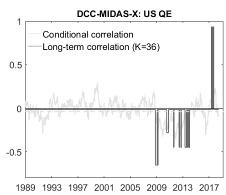


(c) Japan

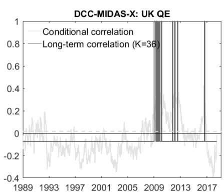


(d) Euro area

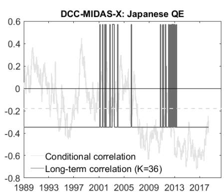
Figure 4.7: Long-term and conditional correlation based on models which include the US QE dummies of Kryzanowski et al. (2017) and a normalisation period dummy in the low-frequency correlation component.



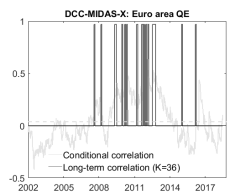
(a) United States



(b) United Kingdom



(c) Japan



(d) Euro area

Figure 4.8: Long-term and conditional correlation based on models which include the QE announcement dummies from Rogers et al. (2014) in the low-frequency correlation component.

Table 4.9 considers the announcement effects of the quantitative easing undertaken in the UK, Japan and the Euro area.[23] The QE announcement effect dummy alone is only statistically significant in Japan, where the positive coefficient implies the effect is opposite to that of the balance sheet data. From Figure 4.8 it is clear that for Japan the timing of the QE announcements do not coincide with the largest negative shifts in the correlation. For the UK the QE dummy seems to capture the same effect as the VXO index while Figure 4.8 suggests QE announcements occur simultaneously with the positive spikes in the correlation, despite their weak statistical significance. For the Euro area the QE announcement dummy is never statistically significant. Overall

[23]For the models also including economic data the QE dummies were mostly not statistically significant, and therefore these results are omitted. Full tables are available upon request.

the results suggest it is the purchases and the balance sheet expansion and not the announcements that are important for the low-frequency correlation between FX returns and equity returns.

Table 4.9: DCC-MIDAS-X model: Quantitative easing dummies

	a	b	θ_c	ω_c	θ_D	$\bar{m}_c$	LLF	AIC	BIC
United Kingdom									
Only QE	0.0135***	0.9829***	-	-	0.4567	-0.0713	-14360.5	3.5609	3.5643
	(0.0034)	(0.0053)			(0.4580)	(0.0545)			
RC	0.0211***	0.9182***	0.8250***	7.0607***	0.1591	-0.0053	-14340.9	3.5565	3.5617
	(0.0053)	(0.0312)	(0.1186)	(1.6434)	(0.1787)	(0.0223)			
CB assets (Δ)	0.0122***	0.9812***	0.8356*	1.1**	0.4678	-0.1221***	-14351.4	3.5591	3.5643
	(0.0039)	(0.0069)	(0.4274)	(0.5111)	(0.9963)	(0.0363)			
VXO index	0.0126***	0.9824***	0.0076	1.1***	1.4915***	-0.2252***	-14356.2	3.5603	3.5655
	(0.0033)	(0.0058)	(0.0051)	(0.2469)	(0.5167)	(0.0868)			
Japan									
Only QE	0.0178***	0.9799***	-	-	1.0118**	-0.3633***	-13164.3	3.4071	3.4107
	(0.0059)	(0.0068)			(0.4607)	(0.0726)			
RC	0.0263***	0.9495***	0.9999***	3.5853***	0.0393	-0.0526	-13143.93	3.4023	3.4077
	(0.0047)	(0.0105)	(0.0779)	(0.8808)	(0.0907)	(0.0352)			
Inflation	0.0176***	0.9800***	0.0951	18.7368	1.3487*	-0.4289***	-13163.1	3.4073	3.4127
	(0.0060)	(0.0072)	(0.0585)	(38.2588)	(0.7334)	(0.0840)			
CB assets (Δ)	0.0181***	0.9793***	-0.1801**	1.1	5.2502	-0.2645***	-13161.5	3.4069	3.4123
	(0.0056)	(0.0075)	(0.0887)	(4.5782)	(70.0615)	(0.0841)			
Euro area									
Only QE	0.0150*	0.9792***	-	-	2.1172	0.0001	-7911.3	3.2453	3.2506
	(0.0082)	(0.0108)			(12.6251)	(0.0924)			
RC	0.0508***	0.8475***	0.8652***	7.3450***	-0.0257	0.0332	-7898.7	3.2410	3.2490
	(0.0168)	(0.0572)	(0.1642)	(2.6051)	(0.1396)	(0.0334)			
CB assets (Δ)	0.0141	0.9786***	0.1318*	1.5980	1.2643	-0.1048	-7907.1	3.2444	3.2524
	(0.0116)	(0.0165)	(0.0707)	(1.2031)	(3.3880)	(0.0974)			

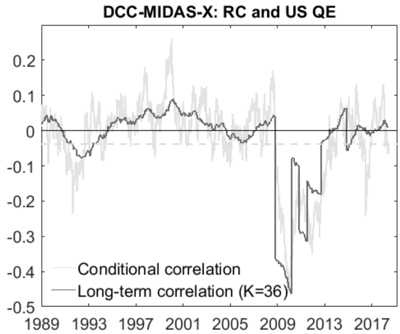
*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. RC denotes realised correlation. The QE dummies follow the dates identified in Rogers et al. (2014). Otherwise, see notes on Table 4.4.*

4.5.6 Multivariate regression model

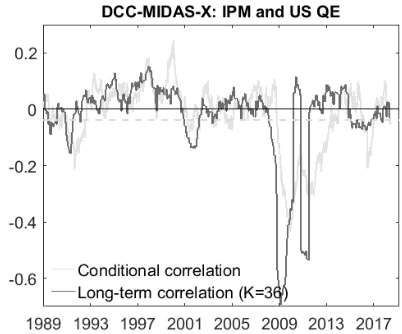
To assess the relative explanatory power of the economic data and as a robustness check, I present results where the (monthly) long-term correlation has been extracted using the DCC-MIDAS-RC model and then, in a second step, a multivariate linear regression model is used to explain the correlation using the economic data.[24] This approach echoes that in Moore and Wang (2014) and Kryzanowski et al. (2017).

I estimate a model which includes all the variables, but multicollinearity issues are resolved by only including one representative of highly correlated variables. The central bank asset data and the monetary base both describe the same phenomenon, and therefore I only include one of them at a time. The interest rate differentials are highly correlated, and hence I only include the long-term yield differential in the regression

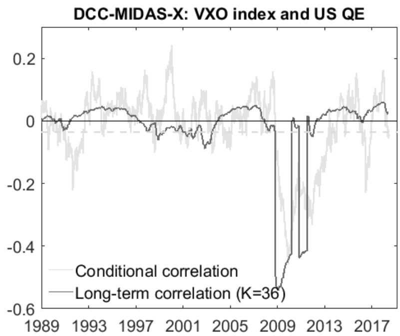
[24] As discussed in Engle et al. (2013) for the case of volatility modelling, using a noisy measure of volatility, such as that obtained from a GARCH-MIDAS model, will lead to bias in the regression parameter estimates of this two-step approach. The DCC-MIDAS-X model circumvents this issue by a single-step estimation procedure.



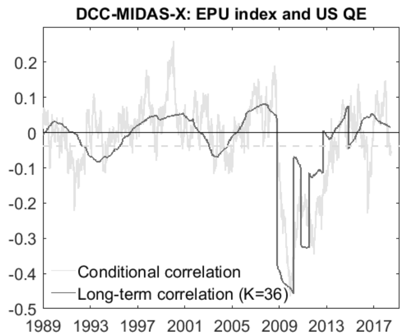
(a) Realised correlation



(b) Industrial prod.



(c) VXO index



(d) EPU index

Figure 4.9: US long-term and conditional correlation when including the US QE dummies of Kryzanowski et al. (2017) and a normalisation dummy as well as economic data in the low-frequency correlation component of a DCC-MIDAS model.

model. For the same reason I also leave out inflation and industrial production growth relative to the US. I use a 36 month moving average of the explanatory data to incorporate a similar lag structure as in the DCC-MIDAS-X models.

The multivariate regression models in Table 4.10 largely confirm the results of the DCC-MIDAS-X models, in the sense that the QE measures (i.e., central bank assets and M0) are mostly highly significant, while the interest rate differential is important for Japan and the Euro area. The Economic Policy Uncertainty index is important for the US, while the VXO index is an important explanatory variables for both the US and Japan. With the exception of the VXO index in the regression for the US, the signs of the (statistically significant) parameter estimates are preserved. In addition, the regression results suggest that inflation is an important variable for especially the UK and Japan. The R^2 are also relatively high for all models, indicating economic variables are important drivers of the correlation.[25]

[25]Results for multivariate regression models including an AR(1) term can be found in Appendix 4.F. The inclusion of an AR term naturally reduces the significance of the economic variables. For the US the VXO

Table 4.10: Multivariate regression models

	US		UK		Japan		Euro area	
	β	β	β	β	β	β	β	β
Constant	-0.0155	-0.0316***	-0.0487***	0.0273	-0.1903***	-0.2503***	-0.0231	-0.0133
	(0.0033)	(0.0042)	(0.0183)	(0.0283)	(0.0456)	(0.0323)	(0.0414)	(0.0248)
IP growth	0.0333***	0.0002	0.1798	0.2315	0.4843	0.5802***	-0.1353	0.0264
	(0.0114)	(0.0127)	(0.1271)	(0.1604)	(0.3039)	(0.2041)	(0.1721)	(0.1000)
Inflation	-0.0039	-0.0117**	0.0989**	0.1156***	0.1867***	0.1802***	0.0487	0.0124
	(0.0047)	(0.0048)	(0.0394)	(0.0442)	(0.0486)	(0.0325)	(0.0412)	(0.0333)
LTY (vs US)	-	-	-0.0460	-0.0737	-0.1582***	-0.0909***	0.2657***	0.2601***
			(0.0529)	(0.0591)	(0.0277)	(0.0324)	(0.0445)	(0.0324)
EPU index	-0.1198***	-0.1199***	-	-	-	-	-	-
	(0.0071)	(0.0060)						
VXO index	0.0236***	0.0217***	0.0394	0.0469	0.1572***	0.1283***	-0.0477	-0.0402
	(0.0059)	(0.0055)	(0.0271)	(0.0312)	(0.0781)	(0.0251)	(0.0869)	(0.0541)
M0	-0.0511***	-	0.5127***	-	-0.1887	-	0.5986***	-
	(0.0192)		(0.0801)		(0.2442)		(0.1036)	
CB assets	-	-0.1660***	-	0.7662***	-	-0.5148***	-	0.7762***
		(0.0639)		(0.0290)		(0.1742)		(0.0982)
R^2	0.64	0.66	0.51	0.46	0.62	0.72	0.64	0.78

*HAC standard errors can be found in parenthesis below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. Model: $Corr_t = \beta \mathbf{X_{t-1}} + \epsilon_t$, where $\mathbf{X}$ is a vector of the explanatory variables, including a constant, and β is the corresponding vector of parameters. The dependent variable is the correlation extracted using the DCC-MIDAS-RC model with 36 lags. All explanatory variables are standardised to have mean 0 and standard deviation 1. A 36 month moving average of each variable is included. LTY denotes the long-term yield. Otherwise, see notes on Table 4.4.*

4.6 Conclusion

This paper studies the link between the low-frequency equity - foreign exchange return correlation and the state of the macro economy and the financial markets using the DCC-MIDAS-X model, which allows directly determining the economic drivers of the long-term correlation. The focus is on variables related to international portfolio rebalancing and on determining how and why the sign of the long-term correlation changes over time, as there are conflicting theoretical predictions for the sign of the correlation. In particular, I consider two hypotheses suggested in the earlier literature, namely flight-to-quality and quantitative easing related search-for-yield, and find support for both channels, especially on the US market. In addition, I investigate whether the correlation is sensitive to standard macroeconomic variables, such as economic growth or the interest rate differential. The interest rate differential is especially important in Japan, most likely reflecting the prominent role of the Yen as a funding currency. The results imply that the negative co-movement of FX and equities strengthens the weaker the carry trade motives. Although there are common elements driving the return correlations, such as interest rate differentials and quantitative easing measures, the results highlight the heterogeneity of the developed markets considered. For instance, while central bank quantitative easing has a negative impact on the correlation in the US and Japan, in line with the portfolio rebalancing channel, in the UK and the Euro area QE has had a positive impact on the correlation. The strong impact of

index retains its importance, while in general the central bank asset data remains at least weakly significant.

the US quantitative easing is concentrated around the onset of the first QE programme in 2008, with QE1 and QE2 being important for the correlation in the US. Overall the results for quantitative easing suggest actual purchases – or being in a QE regime – are more important than announcement effects for the long-term correlation.

This paper complements previous papers studying the causal relationship between stock returns and exchange rates. I also extend the current analysis on correlations, in particular by looking beyond net capital flow data and by using a modelling framework which allows including economic data directly in the correlation model. In addition, I consider several hypotheses related to the portfolio rebalancing channel in a unified framework. The paper could be extended by including a wider variety of markets or asset classes to give a richer picture of the financial market correlations. I concentrated on developed markets because the chosen modelling framework requires long time series, which are still unavailable for many developing markets. Once we understand why and how correlations vary over time it becomes relevant to consider whether this information has implications for economic policy or for investment decisions, and whether the information could be exploited in an out-of-sample forecasting context.

References

Ajayi, R. A. and M. Mougoué (1996). On the dynamic relation between stock prices and exchange rates. *The Journal of Financial Research 19*(2), 193–207.

Altavilla, C., G. Carboni, and R. Motto (2015). Asset purchase programmes and financial markets: lessons from the euro area. *ECB Working paper*.

Asgharian, H., C. Christiansen, and A. J. Hou (2016). Macro-finance determinants of the long-run stock-bond correlation: The DCC-MIDAS specification. *Journal of Financial Econometrics 14*, 617–642.

Bae, J. W. and R. Elkamhi (2016). Global equity correlation in FX carry and momentum trades. *Available at SSRN 2521608*.

Bahmani-Oskooee, M. and A. Sohrabian (1992). Stock prices and the effective exchange rate of the dollar. *Review of Financial Studies 24*(4), 459–464.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Bauwens, L. and E. Otranto (2016). Modeling the dependence of conditional correlations on market volatility. *Journal of Business & Economic Statistics 34*(2), 254–268.

Belke, A., D. Gros, and T. Osowski (2017). The effectiveness of the Fed's quantitative easing policy: New evidence based on international interest rate differentials. *Journal of International Money and Finance 73*, 335–349.

Bernanke, B. S. (2012). Monetary policy since the onset of the crisis: A speech at the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming, August 31, 2012. Speech 645, Board of Governors of the Federal Reserve System.

Branson, W. (1983). *Macroeconomic determinants of real exchange risk, in R.J. Herring (ed): Managing Foreign Exchange Risk.* Cambridge University Press, Cambridge, MA.

Caporale, G. M., J. Hunter, and F. M. Ali (2014). On the linkages between stock prices and exchange rates: Evidence from the banking crisis of 2007–2010. *International Review of Financial Analysis 33*, 87–103.

Cenedese, G., R. Payne, L. Sarno, and G. Valente (2015). What do stock markets tell us about exchange rates? *Review of Finance 20*(3), 1045–1080.

Cho, J.-W., J. H. Choi, T. Kim, and W. Kim (2016). Flight-to-quality and correlation between currency and stock returns. *Journal of Banking & Finance 62*, 191–212.

Colacito, R., R. F. Engle, and E. Ghysels (2011). A component model for dynamic correlations. *Journal of Econometrics 164*, 45–59.

Conrad, C. and K. Loch (2014). Anticipating long-term stock market volatility. *Journal of Applied Econometrics 30*(7), 1090–1114.

Conrad, C., K. Loch, and D. Rittler (2014). On the macroeconomic determinants of long-term volatilities and correlations in US stock and crude oil markets. *Journal of Empirical Finance 29*, 26–40.

Dornbusch, R. and S. Fischer (1980). Exchange rates and the current account. *The American Economic Review 70*(5), 960–971.

Engle, R. F. (2002). Dynamic conditional correlation - a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics 20*, 339–350.

Engle, R. F. (2009). *Anticipating correlations: a new paradigm for risk management*. Princeton: Princeton University Press.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *The Review of Economics and Statistics 95*(3), 776–797.

Engle, R. F. and K. Sheppard (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Technical report, National Bureau of Economic Research.

Frankel, J. A. (1983). *Monetary and portfolio-balance models of exchange rate determination. In: Bhandari J.S., Putnam, B.H. (Eds.), Economic Interdependence and Flexible Exchange Rates.* MIT Press, Cambridge.

Ghalanos, A. (2019). *rmgarch: Multivariate GARCH models.* R package version 1.3-6.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The MIDAS touch: Mixed data sampling regression models. *EFA 2009 Bergen Meetings Paper*.

Ghysels, E., A. Sinko, and R. Valkanov (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Hau, H. and H. Rey (2004). Can portfolio rebalancing explain the dynamics of equity returns, equity flows, and exchange rates? *American Economic Review 94*(2), 126–133.

Hau, H. and H. Rey (2005). Exchange rates, equity prices, and capital flows. *The Review of Financial Studies 19*(1), 273–317.

Kim, K.-H. (2003). Dollar exchange rate and stock price: evidence from multivariate cointegration and error correction model. *Review of Financial Economics 12*.

Kollias, C., N. Mylonidis, and S.-M. Paleologou (2012). The nexus between exchange rates and stock markets: evidence from the euro-dollar rate and composite european stock indices using rolling analysis. *Journal of Economics and Finance 36*(1), 136–147.

Kryzanowski, L., J. Zhang, and R. Zhong (2017). Cross-financial-market correlations and quantitative easing. *Finance Research Letters 20*, 13–21.

Lin, C.-H. (2012). The comovement between exchange rates and stock prices in the Asian emerging markets. *International Review of Economics & Finance 22*(1), 161–172.

Lin, J.-B. and S.-H. Fu (2016). Investigating the dynamic relationships between equity markets and currency markets. *Journal of Business Research 69*, 2193–2198.

Malliaropulos, D. (1998). International stock return differentials and real exchange rate changes. *Journal of International Money and finance 17*(3), 493–511.

Moore, T. and P. Wang (2014). Dynamic linkage between real exchange rates and stock prices: Evidence from developed and emerging Asian markets. *International Review of Economics & Finance 29*, 1–11.

Nieh, C.-C. and C.-F. Lee (2002). Dynamic relationship between stock prices and exchange rates for G-7 countries. *The Quarterly Review of Economics and Finance 41*(4), 477–490.

Pavlova, A. and R. Rigobon (2007). Asset prices and exchange rates. *Review of Financial Studies 20*(4), 1139–1180.

Phylaktis, K. and F. Ravazzolo (2005). Stock prices and exchange rate dynamics. *Journal of International Money and Finance 24*(7), 1031–1053.

Rogers, J. H., C. Scotti, and J. H. Wright (2014). Evaluating asset-market effects of unconventional monetary policy: A multi-country review. *Economic Policy 29*(80), 749–799.

Thornton, D. L. (2014). Has QE been effective? *Economic Synopses* (3).

Tsagkanos, A. and C. Siriopoulos (2013). A long-run relationship between stock price index and exchange rate: A structural nonparametric cointegrating regression approach. *Journal of International Financial Markets, Institutions and Money 25*, 106–118.

Appendices

4.A Data description

Data sources:

- FX data: BIS. Effective exchange rates (EER) refer to the Narrow index with 27 economies and the spot rate to 'home'/USD
- Equity data: Datastream, Alpha Vantage
- Industrial production, Inflation: OECD
- VXO index: Cboe (`http://www.cboe.com/`)
- Short-term interest rates: FRED (Japan, US), OECD (UK, Euro area)
- Long-term interest rates: OECD (UK, Euro area, US), Datastream (Japan)
- Policy rates: Bank of England (UK), FRED (US), Bank of Finland (Euro area), Datastream and BIS (Japan)
- Economic Policy Uncertainty index: FRED
- Central bank balance sheet data (assets) and the monetary base (M0): Federal Reserve (US), Bank of England (UK), European Central Bank (Euro area), Bank of Japan (Japan)

Test for constant conditional correlation

Engle and Sheppard (2001) suggested a test for whether the correlation is constant. Applying this test to the return series reveals that all of the correlations can be considered dynamic. However, for the UK and some of the excess returns the lag length used for the test impacts the conclusion.

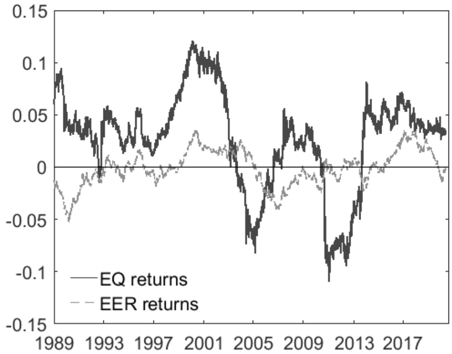
Table 4.11: Test for constant conditional correlation, p-values

Lags:	1	4	20		1	4	20
GBP EER vs EQ	0.0037	0.0000	0.0000	GBP EER vs excess EQ	0.0325	0.0029	0.0002
JPY EER vs EQ	0.0000	0.0000	0.0000	JPY EER vs excess EQ	0.0205	0.0000	0.0000
EUR EER vs EQ	0.0092	0.0000	0.0000	EUR EER vs excess EQ	0.2031	0.1060	0.0000
GBP/USD vs EQ	0.3804	0.0000	0.0000	GBP/USD vs excess EQ	0.0000	0.0000	0.0000
JPY/USD vs EQ	0.0004	0.0000	0.0000	JPY/USD vs excess EQ	0.0002	0.0000	0.0000
EUR/USD vs EQ	0.0001	0.0000	0.0000	EUR/USD vs excess EQ	0.2378	0.0445	0.0000
USD EER vs EQ	0.0847	0.0000	0.0000				

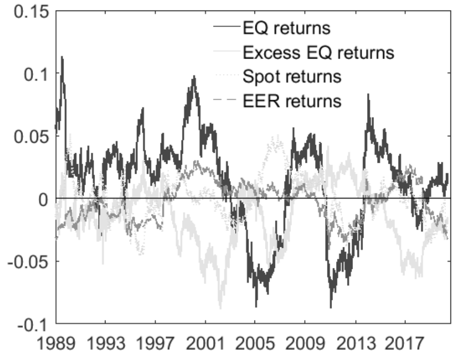
Engle and Sheppard (2001) test for conditional correlation. H0 is that the correlation is constant: low p-values indicate a rejection of the null hypothesis. The test is implemented in R using the rmgarch package presented in Ghalanos (2019).

Figures on the return data

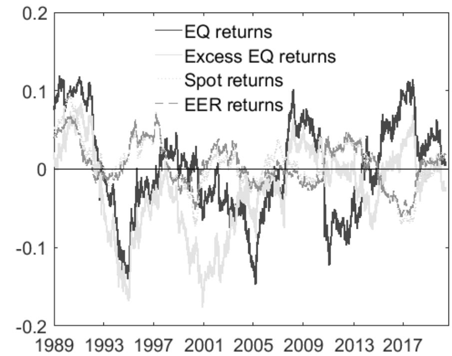
In order to illustrate how the returns vary over time I draw the three year average daily returns in Figure 4.10, calculated over rolling windows. Clearly in the US, the UK and Japan the equity markets share two negative episodes, one around 2005 and another around 2012 (this one is also common to the Euro area). Due to this common global variation excess returns tend to be less negative or even positive during these episodes. However, outside these episodes excess returns tend to be more negative than actual returns. In Japan and the Euro area returns on the spot rate and returns on the effective exchange rate have behaved very similarly.



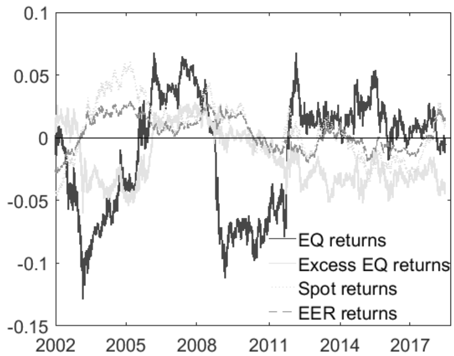
(a) The US



(b) The UK



(c) Japan



(d) Euro area

Figure 4.10: Three year average daily returns and excess returns calculated over rolling windows. EQ refers to equities.

Figures on explanatory variables

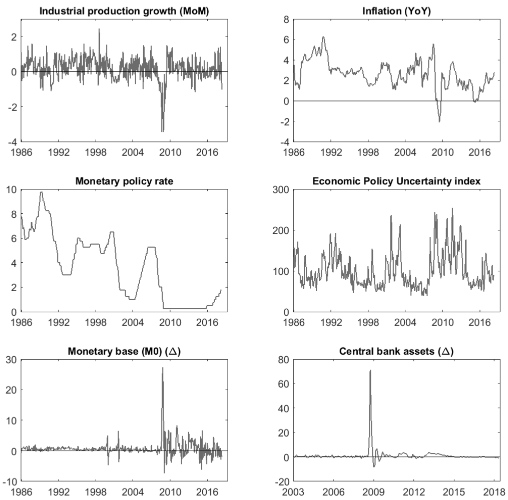


Figure 4.11: Explanatory data for the US.

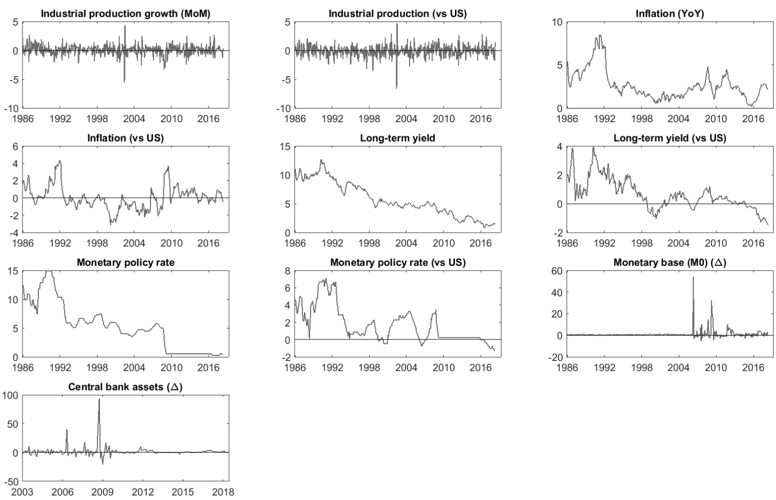


Figure 4.12: Explanatory data for the UK.

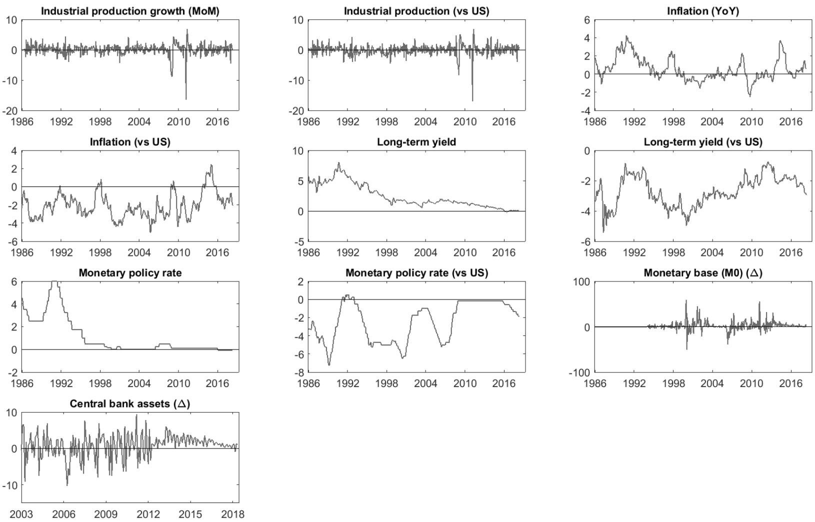


Figure 4.13: Explanatory data for Japan.

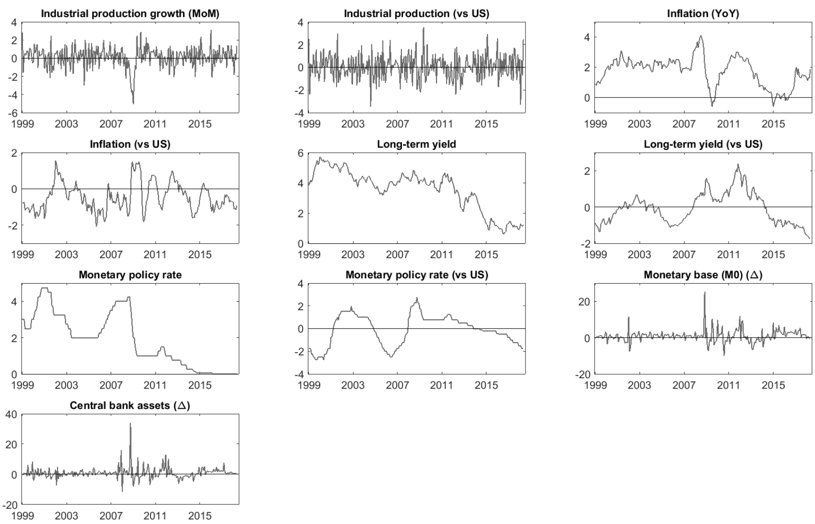


Figure 4.14: Explanatory data for the Euro area.

4.B Dynamic correlations for alternative return data

Figure 4.15 shows the dynamic correlations extracted using the DCC-MIDAS-RC model for excess equity returns and the effective exchange rate for the UK, Japan and the Euro area.

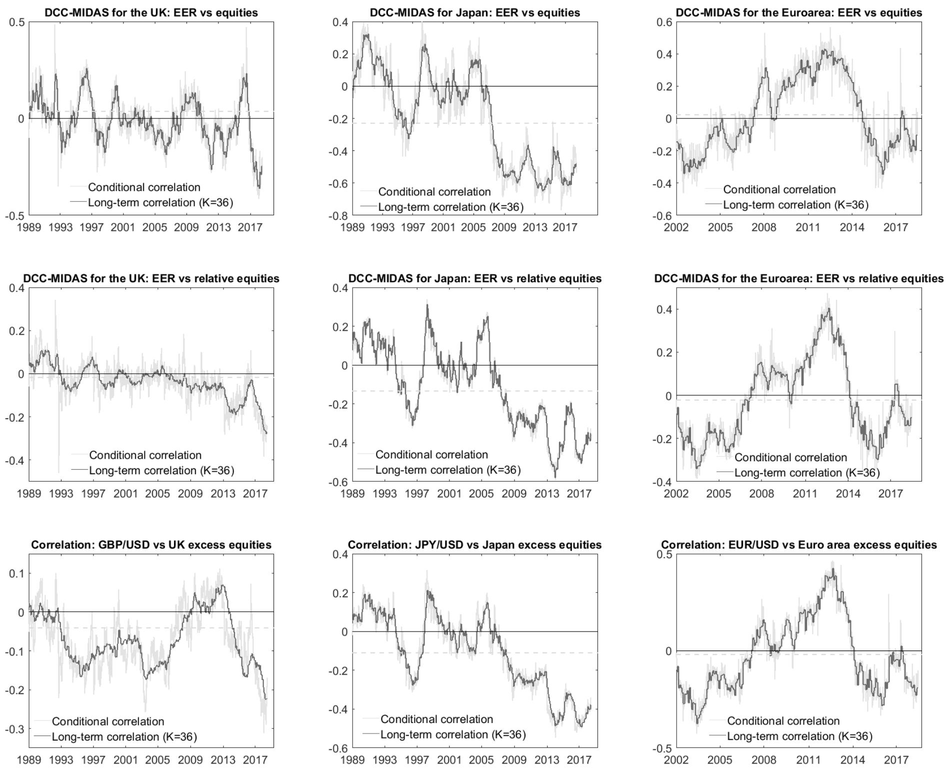


Figure 4.15: Long-term and conditional correlation between equity and FX returns.

4.C GARCH-MIDAS model results: alternative return data

Table 4.12 presents results for estimating the GARCH-MIDAS-RV model using the effective exchange rate and excess equity returns for the UK, Japan and the Euro area. As neither the exchange rates nor excess returns display asymmetric effects, a model without γ is estimated and presented in the table.

Table 4.12: Estimation results of the GARCH-MIDAS-RV model

	μ	α	β	θ_v	ω_2	$\bar{m}_v$	VR
GBP EER returns	0.0042	0.1132***	0.8108***	0.1778***	6.4707***	0.2530***	43.15
	(0.0045)	(0.0337)	(0.0606)	(0.0140)	(1.8858)	(0.0357)	
JPY EER returns	-0.0055	0.0826***	0.8661***	0.1630***	5.8566*	0.4361***	29.46
	(0.0062)	(0.0142)	(0.0432)	(0.0252)	(3.1957)	(0.0794)	
EUR EER returns	0.0055	0.0380***	0.9591***	-	-	0.1991***	-
	(0.0049)	(0.0055)	(0.0059)			(0.0608)	
UK excess eq. ret.	-0.0209**	0.1767***	0.7100***	0.1885***	9.3270***	0.6491***	44.31
	(0.0101)	(0.0238)	(0.0397)	(0.0152)	(2.1990)	(0.0653)	
Japan excess eq. ret.	-0.0066	0.1469***	0.7888***	0.1517***	6.2821***	1.3460***	17.47
	(0.0168)	(0.0227)	(0.0332)	(0.0259)	(1.9585)	(0.2032)	
Euro area excess eq. ret.	-0.0153	0.1463***	0.8021***	0.1863***	8.0550*	0.7839***	33.26
	(0.0143)	(0.0226)	(0.0371)	(0.0214)	(4.7353)	(0.1248)	

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. The MIDAS polynomial: $m_\tau = \bar{m}_v + \theta_v \sum_{k=1}^{K_v} \varphi_k(\omega_v) RV_{\tau-k}$, where RV stands for realised volatility (squared returns). VR stands for the variance ratio, a measure of in-sample fit: $100\frac{Var(log(m_\tau))}{Var(log(m_\tau g_{t,\tau}))}$. eq. ret. denotes equity returns. The models are estimated for the period January 1986-June 2018, with two years of lagged data used for the first period. For the Euro area the sample period starts in January 1999.*

4.D DCC-MIDAS-X model results: excess equity returns

This appendix considers the robustness of the results in Section 4.5.4 to using excess equity returns for the UK, Japan and the Euro area instead of actual returns. Overall the results confirm the conclusions drawn using actual returns. The tables report full estimation results, but to save space only the relevant sub-figures are reported.

Table 4.13: DCC-MIDAS-X model: GBP/USD returns vs. excess UK equity returns

	a	b	θ_c	ω_c	$\bar{m}_c$
Policy rate (vs. US)	0.0074**	0.9840***	0.0039	50	-0.0897***
	(0.0033)	(0.0092)	(0.0124)	(125.29)	(0.0272)
Inflation (YoY)	0.0079**	0.9808***	0.0164	50	-0.1241***
	(0.0032)	(0.0100)	(0.0131)	(56.6140)	(0.0351)
Inflation (vs. US)	0.0075**	0.9832***	0.0142	50	-0.0820***
	(0.0033)	(0.0099)	(0.0188)	(47.6280)	(0.0211)
Long term yield (vs. US)	0.0080**	0.9806***	0.0287	50***	-0.0987***
	(0.0031)	(0.0101)	(0.0213)	(18.5660)	(0.0214)
Monetary base (M0) (Δ)	0.0063**	0.9835***	0.0584***	1.1**	-0.1405***
	(0.0032)	(0.0080)	(0.0209)	(0.5388)	(0.0271)
Central bank assets (Δ)	0.0065*	0.9839***	0.0379**	1.1	-0.1109***
	(0.0034)	(0.0100)	(0.0171)	(1.3183)	(0.0220)
Industrial production growth (MoM)	0.0072**	0.9847***	0.0671	2.5664	-0.0863***
	(0.0032)	(0.0088)	(0.1279)	(2.0299)	(0.0223)
Industrial production growth (vs. US)	0.0071**	0.9838***	0.1162	2.3344	-0.0694***
	(0.0031)	(0.0090)	(0.1095)	(3.2768)	(0.0254)
VXO index	0.0074**	0.9832***	0.0057	1.1	-0.1985**
	(0.0032)	(0.0092)	(0.0045)	(2.4185)	(0.0949)

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. Δ denotes change over period.*

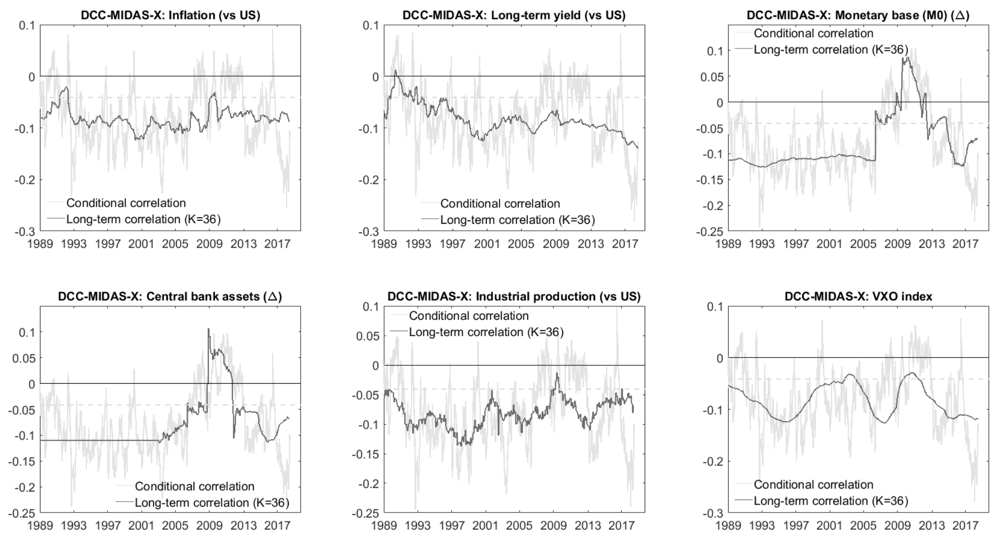


Figure 4.16: Long-term and conditional correlation between the excess UK equity return and the GBP/USD return.

Table 4.14: DCC-MIDAS-X model: JPY/USD return vs. excess Japanese equity return

	a	b	θ_c	ω_c	$\bar{m}_c$
Policy rate (vs. US) (a)	0.0066***	0.9920***	-0.1217***	1.1	-0.5297***
	(0.0011)	(0.0016)	(0.0273)	(1.9969)	(0.0587)
Inflation (YoY)	0.0070***	0.9929***	-26.622	10***	-63.076**
	(0.0021)	(0.0021)	(15.9120)	(0.6674)	(29.732)
Inflation (vs. US)	0.0073***	0.9919***	-0.2418**	1.1	-0.8069***
	(0.0018)	(0.0018)	(0.1035)	(1.1116)	(0.1564)
Long term yield (vs. US)	0.0067***	0.9919***	-0.2770***	1.1	-0.9643***
	(0.0016)	(0.0018)	(0.0487)	(1.3945)	(0.0167)
Monetary base (M0) (Δ) (a)	0.0071***	0.9925***	0.0192	8.0849	-0.5508*
	(0.0008)	(0.0002)	(0.0904)	(30.2290)	(0.3098)
Central bank assets (Δ)	0.0077***	0.9909***	-0.2172***	1.1	-0.1455*
	(0.0019)	(0.0022)	(0.0640)	(1.3482)	(0.0756)
Industrial production growth (MoM)	0.0063***	0.9932***	4.9170	2.7153***	-1.3636
	(0.0023)	(0.0015)	(6.9431)	(0.3315)	(1.6178)
Industrial production growth (vs. US)	0.0071***	0.9926***	1.0661	2.3110***	-0.4015***
	(0.0015)	(0.0013)	(2.3671)	(0.8042)	(0.0647)
VXO index	0.0074***	0.9920***	0.0329	50	-0.9855***
	(0.0019)	(0.0013)	(0.0288)	(50.4910)	(0.3742)

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. (a) = standard errors based on the inverted Hessian, due to problems with the robust standard errors. Δ denotes change over period.*

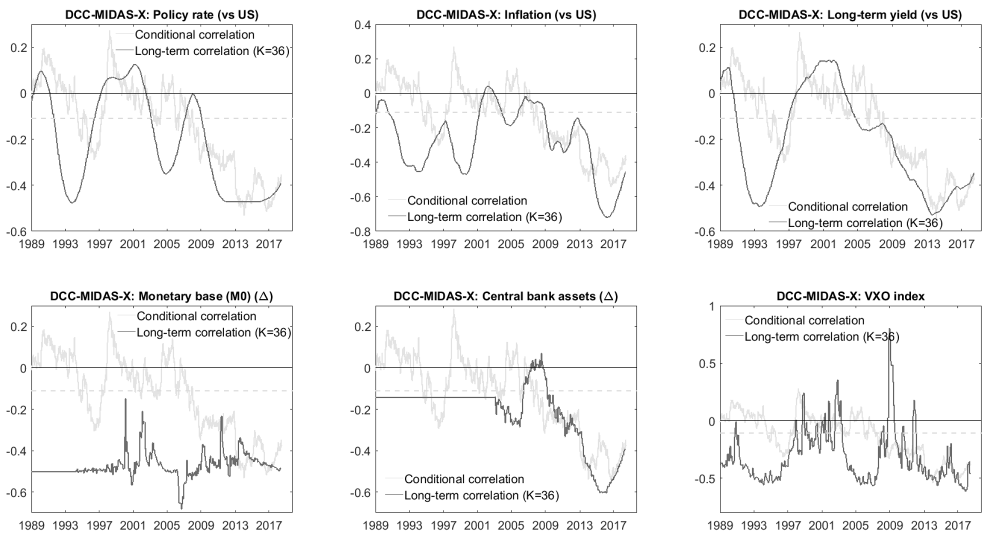


Figure 4.17: Long-term and conditional correlation between the excess Japanese equity return and the JPY/USD return.

Table 4.15: DCC-MIDAS-X model: EUR/USD return vs. excess Euro area equity return

	a	b	θ_c	ω_c	$\bar{m}_c$
Policy rate (vs. US)	0.0102***	0.9871***	0.0634	1.1	0.0246
	(0.0037)	(0.0047)	(0.2162)	(6.9627)	(0.0994)
Inflation (YoY)	0.0103***	0.9873***	-0.0641	1.3203	0.1529
	(0.0036)	(0.0043)	(0.1750)	(1.1332)	(0.3497)
Inflation (vs. US)	0.0102***	0.9876***	-0.0897	43.7950	0.0065
	(0.0039)	(0.0057)	(0.2777)	(268.36)	(0.1584)
Long term yield (vs US)	0.0106***	0.9849***	0.1125	50	0.0198
	(0.0039)	(0.0066)	(0.1184)	(37.6290)	(0.0688)
Monetary base (M0) (Δ)	0.0091	0.9886***	0.1270	10.0130	-0.0930
	(0.0262)	(0.0485)	(1.6308)	(188.09)	(1.6648)
Central bank assets (Δ)	0.0087*	0.9884***	0.1415	9.0740	-0.0958
	(0.0050)	(0.0094)	(0.1947)	(39.1660)	(0.1427)
Industrial production growth (MoM)	0.0106***	0.9866***	-0.1437	35.2080	0.0496
	(0.0039)	(0.0051)	(0.6000)	(32.9970)	(0.1574)
Industrial production growth (vs. US)	0.0095***	0.9885***	0.9342	2.4401	0.0158
	(0.0032)	(0.0041)	(1.0774)	(2.2400)	(0.1123)
VXO index	0.0100**	0.9874***	0.1579	50	-0.2591
	(0.0039)	(0.0054)	(0.0176)	(63.4310)	(0.3376)

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. Δ denotes change over period.*

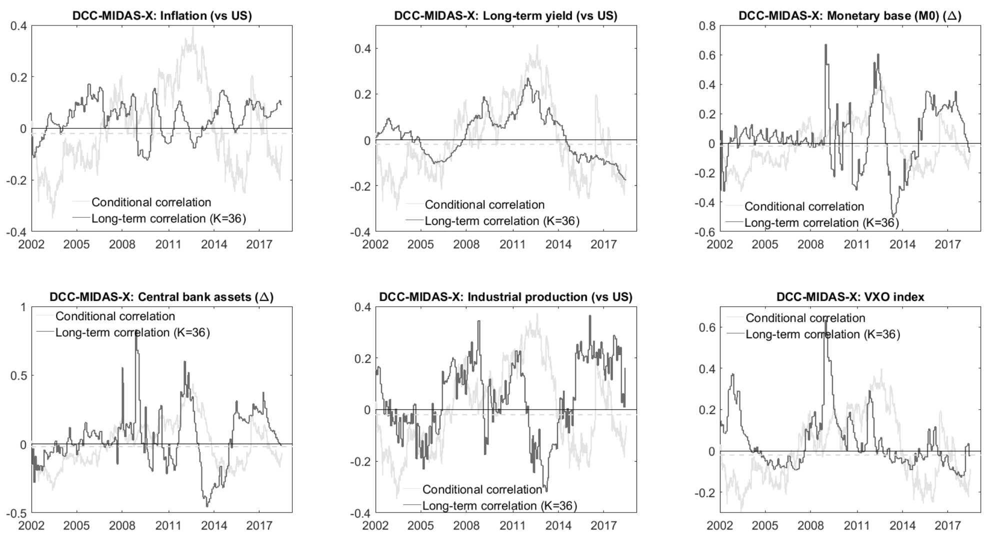


Figure 4.18: Long-term and conditional correlation between the excess Euro area equity return and the EUR/USD return.

4.E DCC-MIDAS-X model results: effective exchange rates

This section considers the robustness of the results in Section 4.5.4 to using the effective exchange rate for the UK, Japan and the Euro area.

Table 4.16: DCC-MIDAS-X model: GBP EER return vs. UK equity market returns

	UK equity return					Excess UK equity return				
	a	b	θ_c	ω_c	$\bar{m}_c$	a	b	θ_c	ω_c	$\bar{m}_c$
Rate (vs. US)	0.0147**	0.9692***	0.0204	50	-0.0405	0.0142**	0.9521***	0.0205	50	-0.0858***
	(0.0061)	(0.0175)	(0.0190)	(31.7315)	(0.0307)	(0.0055)	(0.0185)	(0.0155)	(35.4085)	(0.0221)
CPI (YoY)	0.0140**	0.9722***	0.0264	50	-0.0751	0.0133**	0.9603***	0.1759	50	-0.0965***
	(0.0057)	(0.0150)	(0.0230)	(32.6014)	(0.0554)	(0.0066)	(0.0227)	(0.1525)	(40.6577)	(0.0373)
CPI (vs. US)	0.0141**	0.9720***	0.0331	50**	-0.0072	0.0114	0.9680***	-0.2201	1.0557	-0.0539***
	(0.0055)	(0.0149)	(0.0298)	(22.7577)	(0.0281)	(0.0075)	(0.0275)	(0.2571)	(0.9087)	(0.0202)
LTY (vs. US)	0.0152***	0.9656***	0.0617**	50	-0.0418*	0.0138**	0.9506***	0.0562**	30.9858*	-0.0833***
	(0.0054)	(0.0160)	(0.0290)	(37.2144)	(0.0239)	(0.0055)	(0.0151)	(0.0234)	(17.2935)	(0.0183)
M0 (Δ)	0.0130**	0.9760***	-0.0110	1.5418	0.0021	0.0123*	0.9637***	-0.0220*	5.7195	-0.0313
	(0.0052)	(0.0122)	(0.0296)	(1.0303)	(0.0461)	(0.0070)	(0.0233)	(0.0123)	(5.6190)	(0.0262)
CB assets (Δ)	0.0130**	0.9761***	-0.0073	4.9503	-0.0035	0.0126*	0.9620***	-0.0173*	3.4657	-0.0394*
	(0.0052)	(0.0121)	(0.0148)	(6.6106)	(0.0334)	(0.0068)	(0.0243)	(0.0107)	(2.9348)	(0.0235)
IP (MoM)	0.0127**	0.9761***	-0.0941	7.0974***	-0.0065	0.0122**	0.9592***	0.2303*	1.1693***	-0.0624***
	(0.0050)	(0.0115)	(0.1098)	(2.7427)	(0.0290)	(0.0062)	(0.0257)	(0.1181)	(0.3461)	(0.0196)
IP (vs. US)	0.0126***	0.9755***	0.1273	36.1253	0.0070	0.0122	0.9666***	0.0020	4.9847	-0.0533***
	(0.0046)	(0.0112)	(0.0815)	(27.8147)	(0.0303)	(0.0075)	(0.0259)	(0.0188)	(15.0859)	(0.0206)
VXO index	0.0129**	0.9753***	0.0044	13.2586	-0.0972	0.0122*	0.9642***	0.0042	34.3663	-0.1358**
	(0.0050)	(0.0123)	(0.0037)	(8.5501)	(0.0790)	(0.0067)	(0.0233)	(0.0028)	(38.8972)	(0.0579)

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. IP denotes industrial production growth, CB central bank, Rate the policy rate, LTY the long-term yield, and M0 is the monetary base.*

Table 4.17: DCC-MIDAS-X model: JPY EER return vs. Japanese equity market returns

	Japanese equity return					Excess Japanese equity return				
	a	b	θ_c	ω_c	$\bar{m}_c$	a	b	θ_c	ω_c	$\bar{m}_c$
Rate (vs. US)	0.0152***	0.9824***	-0.1048***	1.001	-0.5542***	0.0077***	0.9907***	-0.1128**	1.001	-0.4768***
	(0.0046)	(0.0058)	(0.0377)	(2.6412)	(0.0827)	(0.0017)	(0.0032)	(0.0565)	(5.2760)	(0.0764)
CPI (YoY)	0.0145***	0.9841***	0.0475	40.4219	-0.3586***	0.0079***	0.9913***	0.2037	50	-0.3990***
	(0.0053)	(0.0062)	(0.0730)	(87.7158)	(0.1107)	(0.0017)	(0.0020)	(0.1602)	(16.6825)	(0.1365)
CPI (vs. US)	0.0152***	0.9829***	-0.1747**	1.001	-0.6761***	0.0083***	0.9906***	-0.2350***	1.001	-0.7228***
	(0.0051)	(0.0062)	(0.0718)	(0.8519)	(0.1813)	(0.0017)	(0.0019)	(0.0750)	(0.8872)	(0.1519)
LTY (vs. US)	0.0155***	0.9816***	-0.2788***	1.001	-1.0315***	0.0078***	0.9904***	-0.2550***	1.001	-0.8728
	(0.0056)	(0.0077)	(0.0630)	(1.9847)	(0.1606)	(0.0017)	(0.0027)	(0.0668)	(2.2669)	(0.1290)
M0 (Δ)	0.0144*	0.9843***	0.0180	1.7030	-0.3733	0.0082***	0.9914***	0.6223	7.5907	-0.4715**
	(0.0074)	(0.0088)	(0.0674)	(1.6670)	(0.2853)	(0.0015)	(0.0014)	(0.6790)	(12.4563)	(0.2287)
CB assets (Δ)	0.0153**	0.9826***	-0.1638	1.001	-0.2344***	0.0086***	0.9899***	-2.1530***	1.001	-0.1173
	(0.0061)	(0.0089)	(0.5038)	(24.1383)	(0.0654)	(0.0017)	(0.0019)	(0.6317)	(0.9766)	(0.0811)
IP (MoM)	0.0216	0.9864***	1.4588	2.6978***	-0.5325	0.0084***	0.9907***	0.0580	6.3453	-0.2542***
	(0.0098)	(0.0113)	(2.5567)	(0.4819)	(0.5523)	(0.0018)	(0.0022)	(0.1270)	(5.8641)	(0.0941)
IP (vs. US)	0.0146***	0.9840***	-0.0533	6.8837***	-0.3402***	0.0082***	0.9913***	0.6471	31.7160	-0.3882
	(0.0050)	(0.0058)	(0.1928)	(2.3136)	(0.0967)	(0.0016)	(0.0016)	(0.7717)	(44.1344)	(0.2592)
VXO index	0.0148***	0.9836***	0.0168	50***	-0.6561**	0.0083***	0.9909***	0.0411**	1.0831	-1.0900***
	(0.0055)	(0.0067)	(0.0102)	(15.9981)	(0.2608)	(0.0017)	(0.0019)	(0.0175)	(1.0391)	(0.3217)

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. IP denotes industrial production growth, CB central bank, Rate the policy rate, LTY the long-term yield, and M0 is the monetary base.*

Table 4.18: DCC-MIDAS-X model: EUR EER return vs. Euro area equity market returns

	Euro area equity return					Excess Euro area equity return				
	a	b	θ_c	ω_c	$\bar{m}_c$	a	b	θ_c	ω_c	$\bar{m}_c$
Rate (vs. US)	0.0117***	0.9862***	-0.0067	4.9221	0.0922	0.0087***	0.9896***	-0.0906	10.1828	0.0209
	(0.0038)	(0.0051)	(0.1468)	(47.8514)	(0.1225)	(0.0020)	(0.0025)	(0.1454)	(10.6491)	(0.1197)
CPI (YoY)	0.0117***	0.9861***	0.0265	9.8179	0.0444	0.0088***	0.9896***	-1.8382	1.3822	0.3398
	(0.0040)	(0.0051)	(0.1418)	(63.2172)	(0.2763)	(0.0020)	(0.0022)	(2.0596)	(0.9249)	(0.4089)
CPI(vs. US)	0.0118***	0.9861***	0.0148	7.0826	0.0954	0.0087***	0.9899***	-0.2316	50	-0.0792
	(0.0040)	(0.0053)	(0.2379)	(6.6387)	(0.1350)	(0.0022)	(0.0042)	(0.7960)	(158.9862)	(0.3277)
LTY (vs. US)	0.0125***	0.9835***	0.1233	50	0.0691	0.0086	0.9901***	-0.1579	1.7871	0.0283
	(0.0045)	(0.0079)	(0.1433)	(39.6075)	(0.0795)	(0.0156)	(0.0494)	(9.4446)	(8.6370)	(0.6035)
M0 (Δ)	0.0099	0.9882***	0.2112	2.0108	-0.0956	0.0079***	0.9904***	0.1598	4.7451	-0.1443
	(0.0110)	(0.0161)	(0.5231)	(5.0204)	(0.5638)	(0.0024)	(0.0039)	(0.2114)	(11.6105)	(0.2609)
CB assets (Δ)	0.0098	0.9872***	0.2235	1.6556	-0.1105	0.0081***	0.9882***	0.1969*	1.001	-0.1624*
	(0.0088)	(0.0152)	(0.1517)	(3.7846)	(0.1275)	(0.0028)	(0.0035)	(0.1181)	(1.4485)	(0.0969)
IP (MoM)	0.0117***	0.9860***	-0.1940	22.9402	0.1057	0.0063	0.9930***	64.2460	9.0098	-1.2985
	(0.0038)	(0.0055)	(1.2467)	(29.2721)	(0.2475)	(0.0101)	(0.0105)	(677.4559)	(53.9937)	(14.2256)
IP (vs. US)	0.0105***	0.9879***	1.5177	1.001	0.0650	0.0079***	0.9908***	17.0941	2.9924	-0.0386
	(0.0036)	(0.0047)	(3.4786)	(4.1396)	(0.1477)	(0.0019)	(0.0022)	(16.4313)	(2.8365)	(0.1372)
VXO index	0.0113***	0.9865***	0.0131	50	-0.1613	0.0089***	0.9892***	0.0064	1.001	-0.1248
	(0.0037)	(0.0049)	(0.0208)	(68.9140)	(0.4005)	(0.0020)	(0.0025)	(0.0159)	(4.6208)	(0.3457)

*Bollerslev-Wooldridge robust standard errors are reported below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. $K = 36$. IP denotes industrial production growth, CB central bank, Rate the policy rate, LTY the long-term yield, and M0 is the monetary base.*

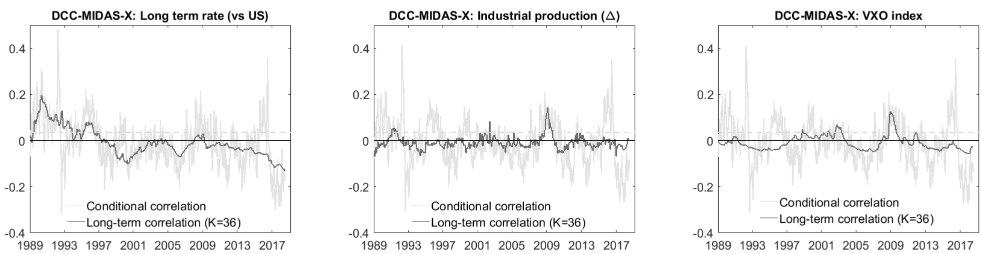


Figure 4.19: Correlation between the UK equity return and the GBP EER return.

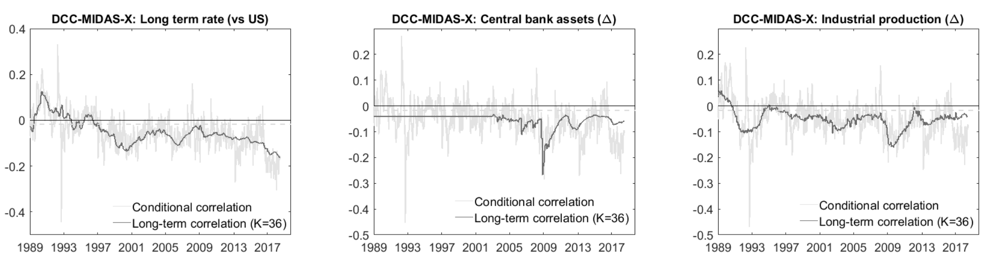


Figure 4.20: Correlation between the excess UK equity return and the GBP EER return.

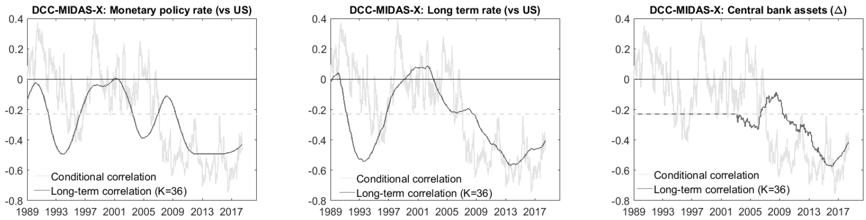


Figure 4.21: Correlation between Japanese equity return and JPY EER return.

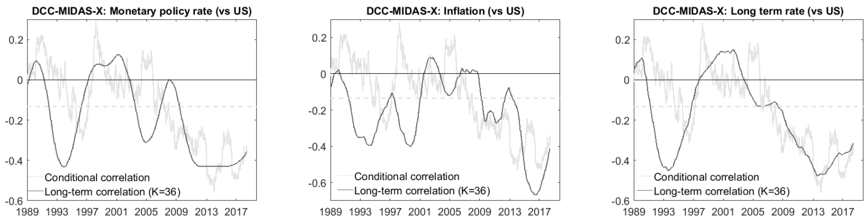


Figure 4.22: Correlation between excess Japanese equity return and JPY EER return.

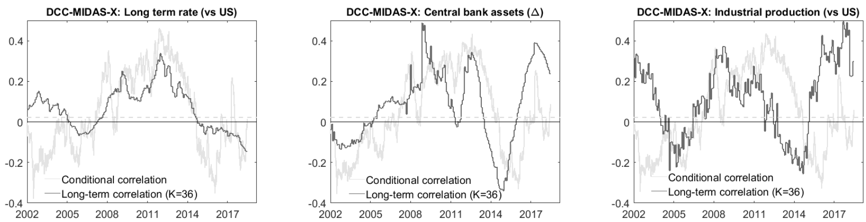


Figure 4.23: Correlation between Euro area equity return and EUR EER return.

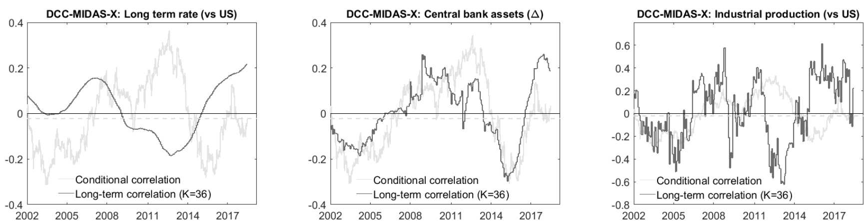


Figure 4.24: Correlation between excess Euro area equity return and EUR EER return.

4.F Multivariate regression model: including an AR term

This appendix presents results for the multivariate regression models which include an autoregressive term. As expected, the AR term is highly significant and reduces the significance of the other variables. Only the central bank balance sheet data consistently retains its importance. For the US the VXO index is negative and highly significant.

Table 4.19: Multivariate regression model with an AR term

	US		UK		Japan		Euro area	
	β	β	β	β	β	β	β	β
Constant	-0.0159***	-0.0169***	-0.0532***	-0.0464***	-0.1924***	-0.1969***	0.0266***	0.0261***
	(0.0006)	(0.0007)	(0.0025)	(0.0035)	(0.0027)	(0.0034)	(0.0039)	(0.0040)
$Corr_{t-1}$	0.1020***	0.1017***	0.1683***	0.1695***	0.2779***	0.2722***	0.2295***	0.2236***
	(0.0009)	(0.0010)	(0.0044)	(0.0042)	(0.0040)	(0.0042)	(0.0057)	(0.0076)
IP	0.0032	0.0022	0.0301	0.0345*	-0.0029	0.0145	-0.0155	-0.0065
	(0.0021)	(0.0023)	(0.0184)	(0.0200)	(0.0149)	(0.0163)	(0.0216)	(0.0212)
Inflation	-0.0004	-0.0008	0.0041	0.0048	-0.0096**	-0.0055	0.0025	0.0004
	(0.0008)	(0.0008)	(0.0065)	(0.0065)	(0.0046)	(0.0040)	(0.0058)	(0.0059)
LTY (vs. US)	-	-	0.0008	-0.0011	-0.0034	-0.0019	0.0029	0.0087
			(0.0070)	(0.0069)	(0.0046)	(0.0043)	(0.0083)	(0.0095)
EPU index	0.0028*	0.0018	-	-	-	-	-	-
	(0.0015)	(0.0014)						
VXO index	-0.0042***	-0.0044***	0.0057	0.0061	0.0054	0.0053	0.0055	0.0047
	(0.0011)	(0.0011)	(0.0044)	(0.0044)	(0.0054)	(0.0048)	(0.0071)	(0.0072)
Monetary base	-0.0060*	-	0.0481***	-	-0.0168	-	0.0388	-
	(0.0032)		(0.0181)		(0.0112)		(0.0239)	
Central bank assets	-	-0.0096*	-	0.0689**	-	-0.0412**	-	0.0615*
		(0.0050)		(0.0270)		(0.0162)		(0.0315)
R^2	0.99	0.99	0.94	0.94	0.99	0.99	0.99	0.98

*HAC standard errors can be found in parenthesis below the parameter estimates. *, ** and *** indicate significance at the 10%, 5%, and 1% level, respectively. Model: $Corr_t = \beta \mathbf{X_{t-1}} + \epsilon_t$, where $\mathbf{X}$ is a vector of the explanatory variables, including a constant and an AR(1) term, and β is the corresponding vector of parameters. The data is standardised to have mean 0 and standard deviation 1. The dependent variable is the correlation extracted using the DCC-MIDAS model with 36 lags. A 36 month moving average of each variable is included. IP denotes industrial production growth and LTY is the long term yield.*